LACIE-00630
JSC-13743
79-10044
TM-79911

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)

NASA NOAA USDA

LACIE PHASE III

ACCURACY ASSESSMENT PLAN

(E79-10044) LARGE AREA CROP INVENTORY EXPERIMENT (LACIE). LACIE PHASE 3 ACCURACY ASSESSMENT PLAN (NASA) 137 p HC A07/MF A01 CSCL 02C
N79-13451
Unclas
G3/43 00044

National Aeronautics and
Space Administration

Lyndon B. Johnson Space Center
Houston, Texas 77058

September 1978

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)

PHASE III ACCURACY ASSESSMENT PLAN

APPROVED BY

D. E. Pitts, Manager
LACIE Accuracy Assessment, NASA/JSC

J. D. Erickson, Chief
RTEB, NASA/JSC

J. L. Dragg, Chief
ASVB, NASA/JSC

F. G. Hall, LACIE Project Scientist
NASA/JSC

J. Hill, Assistant Deputy Project Manager
for LACIE, NOAA

J. Murphy, Assistant Deputy Project Manager
for LACIE, USDA

R. B. Erb, Chief
LACIE Project Office, NASA/JSC

R. B. MacDonald, Chief
Earth Observations Division, NASA/JSC

PRECEDING PAGE BLANK NOT FILMED

CONTENTS

Section | Page

Section Page

Appendix | Page

TABLES

FIGURES

PRECEDING PAGE BLANK NOT FILMED

1. INTRODUCTION

The Large Area Crop Inventory Experiment (LACIE) is an interagency endeavor of the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), and the U.S. Department of Agriculture (USDA). Its purposes are to test and demonstrate the accuracy and economic importance of utilizing remotely sensed data from the Land Satellite (Landsat) in conjunction with climatological and conventional agriculture data to produce timely estimates of the production of a major world crop prior to harvest.

The LACIE Accuracy Assessment (AA) effort is designed to check the accuracy of LACIE products (estimates of wheat production, area, and yield) throughout the growing season, to determine if the operational procedures are sufficient to satisfy LACIE project goals and objectives, and to identify problem areas in the estimation process.

Most of the AA studies are conducted in the U.S. Great Plains (USGP) region and Canada since these regions have the most ground-truth data available. However, the AA studies in these regions are also designed to promote the development of LACIE procedures which can be used to obtain accurate estimates for other parts of the world.

1.1 PHASE III AA OBJECTIVES

The following are the specific objectives for AA in Phase III.

a. To make comparisons throughout the growing season between the LACIE and the USDA Statistical Reporting Service (SRS) reference standard estimates of wheat production, area, and yield for the USGP region.

b. To determine whether the LACIE production estimates are meeting the 90-90 project goal (to be within 10 percent of the true value with a confidence of 90 percent).

c. To conduct investigations of the error sources in the LACIE estimates and, where possible, to relate these error sources to causal elements in the LACIE estimation processes.

d. To assess the accuracy of the LACIE estimates for the U.S.S.R. to the extent permitted by the available data. This will include comparisons throughout the growing season between the LACIE estimates and the USDA Foreign Agricultural Service (FAS) estimates for the U.S.S.R.

e. To investigate the accuracy of LACIE proportion estimates in the U.S. Great Plains and in selected regions of Canada, using blind site data.

This document was prepared by Lockheed Electronics Company, Inc. (LEC), Systems and Services Division, Houston, Texas, under contract NAS 9-15200 for the Earth Observations Division, Space and Life Sciences Directorate, Lyndon B. Johnson Space Center (JSC) of NASA. Inputs were received from NASA, USDA, and NOAA personnel.

2. BACKGROUND

In the LACIE project, remote sensing technology is used in conjunction with meteorological and conventional agricultural information to examine three global crop seasons, each of which is designated as a LACIE phase.

Phase I, which began in January 1975, was devoted primarily to identifying and estimating wheat acreage in the USGP states of Colorado, Kansas, Minnesota, Montana, Nebraska, North Dakota, Oklahoma, South Dakota, and Texas. Classification analyses were conducted in other selected areas, and yield model development and yield feasibility determinations were made over selected regions in the United States. Data from the USDA/SRS were compared with LACIE estimates to determine the accuracy of LACIE performance.

Phase I AA activities were initiated in July 1975, and tests for the accuracy of wheat acreage estimates were conducted using segments for which ground-truth data were available. Initially, statistical tests and comparisons of LACIE estimates with ground-truth data were made using data from 27 intensive test sites (ITS's) in eight states; then, to test a greater number of acquisitions in a more concentrated area, ground observations of harvested small grains were obtained from 30 LACIE operational segments in North Dakota and Montana. The identity of these sites was withheld from the Classification and Mensuration Subsystem (CAMS) analysts so that they would process them as ordinary segments. For this reason, they were called blind sites. After the data from the blind sites were processed by CAMS, the AA Team (appendix A) compared the results of the various sampling and classification procedures used. Approximately 340 special classification runs were conducted to support Phase I AA activities.

In Phase II, which began in October 1975 and covered the 1975-76 growing season, emphasis was on the evaluation of LACIE procedures over the USGP region and on the development of AA methodology.

LACIE acreage, yield, and production estimates for spring and winter wheat were compared with the corresponding USDA/SRS estimates. This was done for the USGP region, for various subregions of the USGP (see section 4.2), and for each state in the USGP. Detailed error source investigations were made by comparing LACIE proportion estimates with ground-truth proportions for over 150 blind sites and 27 ITS's. Estimates of the production, coefficient of variation (CV), and bias were used to evaluate the 90-90 criterion at the USGP level; and a sensitivity analysis was performed to determine the effect of various errors on the LACIE production estimate. In the foreign area, 10 ITS's in Canada were studied and evaluated.

In Phase III, which began in October 1976, the emphasis of AA continues to be the detailed evaluation of LACIE estimates and procedures over the nine-state USGP "yardstick" wheat region. Investigations will also be carried out in the U.S.S.R. and Canada, as mentioned in section 1.1.

3. SCOPE

To accomplish its objectives, AA requires the multiagency collection of various types of data to support the Phase III evaluations. This includes:

a. Aircraft photography over all U.S. blind sites and ITS's and Canadian test sites and ITS's in order to prepare maps for land-use annotations by USDA Agricultural Stabilization and Conservation Service (ASCS) personnel.

b. Land-use annotations from ground observations by USDA/ASCS personnel over all U.S. blind sites and ITS's and Canadian test sites and ITS's.

c. Special ground observances of plant height and ground cover data every 18 days over 15 fields from each U.S. blind site.

d. LACIE imagery, interpretation, and classification data over all U.S. blind sites and ITS's and Canadian test sites and ITS's.

e. Monthly wheat production, area, and yield estimates over the USGP from LACIE and USDA/SRS.

f. Monthly wheat production, area, and yield estimates over regions of the U.S.S.R. and Canada from LACIE and USDA/FAS.

The following is the distribution of blind sites, test sites, and ITS's in the United States and Canada.

a. 143 winter wheat blind sites in the USGP

b. 68 spring wheat blind sites in the USGP

c. 30 spring wheat test sites in Saskatchewan (Canada)

d. 24 ITS's in the USGP

e. 10 ITS's in the Canadian spring wheat region

Specific details of the scope and data requirements for Phase III AA are presented in table 3-1.

Listings and locations of ITS's and distributions of blind sites by state are presented in appendix B.

TABLE 3-1.— SUMMARY OF PHASE III ACCURACY ASSESSMENT DATA REQUIREMENTS

DATA REQUIRED / REGIONS OR SITES	CURRENT YEAR PREDICTION DATA			GROUND TRUTH DATA							LACIE CLASS IMAGERY		LACIE CLASSIFICATION DATA																					
													ERIPS DATA																	IMAGE 100 DATA				
													SEGMENT LEVEL DATA											PIXEL LEVEL DATA								AI DOT LABELS		
																								AI DOT LABELS										
	LACIE CMR P, A, Y ESTIMATES & VARIANCE	USDA/SRS P, A, Y ESTIMATES & VARIANCE	USDA/FAS P, A, Y ESTIMATES & VARIANCE	AIRCRAFT PHOTOGRAPHY	ITS FALL INVENTORY	BLIND SITE FALL INVENTORY	FULL 18 DAY ITS PERIODIC INVENTORY	BS PERIODIC (15 FIELD INVENTORY) PLANT HT GROUND COVER	ITS SPRING INVENTORY	BLIND SITE AT HARVEST INVENTORY	LACIE IMAGERY PFC CONTROL POINTS	LANDSAT TAPE PIXEL DATA	NO TRAINING FIELDS SELECTED	NO DOTS LABELED	CAMS EVALUATION CODE	NO ACQUISITIONS PER SITE	TIME OF ACQUISITIONS	AI IDENTIFICATION	AI ESTIMATES OF BIOSTAGE	P_W, P_{SG}	CORRECTED P_W	TRAINING FIELD VERTICES CARDS	P_{CC} OMISSION P_{CC} COMMISSION	PRECLASS	POSTCLASS	CLASS DOT LABELS	CLUSTER MAPS	CLASS. MAPS	CLASS. SUMMARY	P_W, P_{SG}	CORRECTED P_W CORRECTED P_{SG}	INITIAL IMAGE-100 DOT LABELS	REEVALUATED IMAGE 100 DOT LABELS	RELABLED CLUSTERS
U.S. GREAT PLAINS (9 STATES) FOR EACH STATE THROUGHOUT SEASON	X	X																																
U.S.S.R.	X		X																															
CANADIAN TEST SITES (30)				X						X	X	X		X	X	X	X	X	X	X	X			X	X	X	X	X	X	X	X	X	X	X
U.S. BLIND SITES (211) — WINTER WHEAT SITES (143) — EARLY SEASON WINTER WHEAT SITES (55)				X		X PLUS 15 FIELDS • PT HT • GD.CR.		X STARTS IN MAY		X	X	X	X	X	X	X	X	X	X	X	X	X (PARTIAL)		X	X	X	X	X	X					
U.S. BLIND SITES (211) — WINTER WHEAT SITES (143) — OTHER WINTER WHEAT SITES (88)				X				X STARTS IN MAY		X	X	X	X (PARTIAL)	X	X	X	X	X	X	X	X	X (PARTIAL)		X	X	X	X	X	X					
U.S. BLIND SITES (211) — SPRING WHEAT SITES (68)				X				X STARTS IN MAY		X	X	X		X	X	X	X	X	X	X	X			X	X	X	X	X	X					
U.S. ITS'S (24)				X	X		X STARTS MAR WW MAY SW		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CANADIAN ITS'S (10)				X	X		X		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

PRECEDING PAGE BLANK NOT FILMED

4. GENERAL TECHNICAL APPROACH

Three groups of activities are required to implement the LACIE Phase III AA plan and satisfy its objectives:

a. Planning and data acquisition

b. Data analysis and evaluation

c. Reporting

A detailed flow diagram of the Phase III AA program and related LACIE operational activities is presented in figure 4-1. Specific descriptions of the AA tasks associated with each group of activities are presented in section 6.

4.1 PLANNING AND DATA ACQUISITION

The first group of Phase III AA activities consists of identifying AA data requirements and monitoring the acquisition of the data. Most of the data will be acquired by LACIE operations (NASA), the USDA, and NOAA; but coordination is required by AA personnel to ensure that the data collected are adequate to support the AA program.

Part of this monitoring activity is conducted by members of the AA Team (appendix A). In particular, they select the blind sites in the United States and Canada and coordinate action for acquiring ground-truth data from the blind sites and ITS's in the United States and Canada.

The specific tasks to implement the above activity are identified as tasks 1 through 8 in figure 4-1. Specific descriptions of these tasks are provided in section 6.1.

PHASE III ACCURACY ASSESSMENT PROGRAM

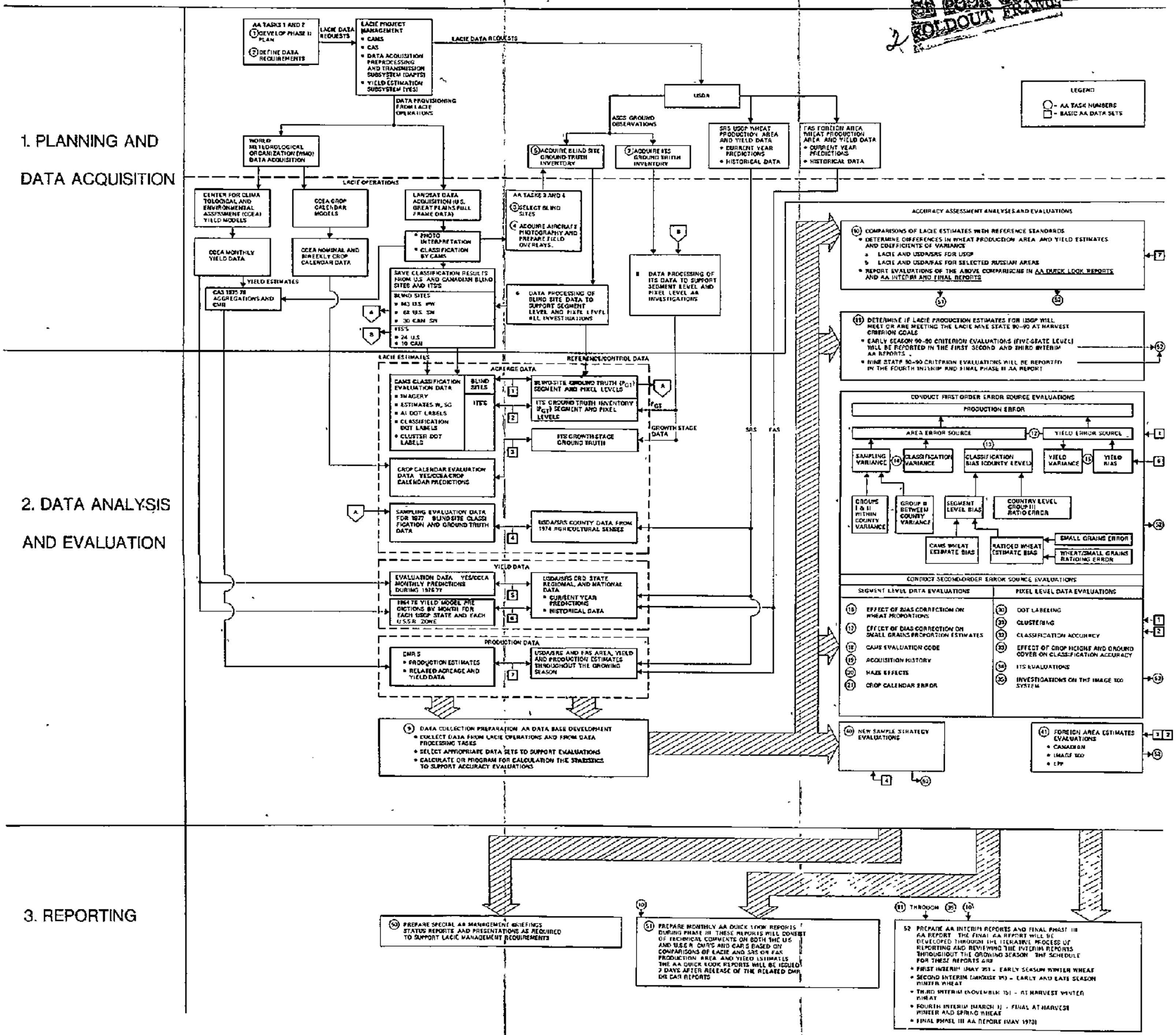


Figure 4-1.— Schematic diagram of LACIE Phase III Accuracy Assessment.

4.2 DATA ANALYSIS AND EVALUATION

The second group of Phase III AA activities involves the analysis of the data collected during the Phase III data acquisition activities. Various subsets of the Phase III data are prepared to support the analyses performed by AA. These basic data sets are identified by the numbers in the square symbols in figure 4-1. Similar numbers on the right side of the diagram indicate the specific analysis to which each data set is applied.

The Phase III AA evaluations are carried out at the country level for the U.S.S.R., at selected sites in the Province of Saskatchewan in Canada, and in the nine-state USGP "yardstick" region in the United States. Within this region, evaluations are usually made for each individual state and for the following smaller regions.

a. The U.S. southern Great Plains region (USSGP) — This region consists of Colorado, Kansas, Nebraska, Oklahoma, and Texas. These states have winter wheat only and therefore could also be called the "winter wheat states." LACIE estimates of wheat production are available for the USSGP from February through October.

b. The spring wheat states [(SW states), Minnesota and North Dakota] — These states have spring wheat only. LACIE estimates of wheat production are available from August through October.

c. The mixed wheat states [(MW states), Montana and South Dakota] — These states have both spring and winter wheat. LACIE estimates of wheat production are available from August through October for spring wheat and from February through October for winter wheat.

d. The U.S. northern Great Plains region (USNGP) — This region consists of the two spring wheat states and the two mixed

e. The USGP region — This region consists of the nine states of the USSGP and the USNGP.

To determine the magnitude and components of error in LACIE estimates and to ascertain whether or not the LACIE production estimates are satisfying the 90-90 criterion, AA personnel do the following:

a. Determine the relative differences[1] between LACIE and USDA/SRS estimates of wheat production, area, and yield over the various regions of the USGP "yardstick" region and between LACIE and USDA/FAS estimates of wheat production, area, and yield in the U.S.S.R. In addition, significance tests are made to compare these estimates. These data are reported throughout the Phase III growing season in the various AA quick-look reports, in the interim AA reports, and in the final AA report.

b. Determine if LACIE production estimates are meeting the LACIE project 90-90 goal. These evaluations are carried out at the country level for the U.S.S.R. and at the USGP level for the United States. In the United States, early season evaluations are based on five-state or seven-state winter wheat data projected to the nine-state level and are reported in the first and second AA interim reports. Nine-state 90-90 criterion evaluations are made when the production estimate and standard error become available for the USGP. They are reported in the third and fourth interim reports.

c. Conduct detailed investigations of error sources within LACIE production estimates for the USGP, which shall consist of the following.

---

[1]Relative difference = $\frac{\text{LACIE - SRS}}{\text{LACIE}}$.

- First-Order Error Source Investigations. These are studies of those errors contributing to LACIE production estimates which can be estimated using LACIE estimates, reference estimates (USDA/SRS), and historical and blind site data. The effect of each error component is assessed by estimating the error in LACIE production estimates caused by removing that error. The first-order error components to be evaluated are:

  a. Yield Error Source Estimation

     (1) Yield Bias
     (2) Yield Variance

  b. Area Error Source Estimation

     (1) Sampling Variance Estimation
     (2) Classification Variance Estimation
     (3) Classification Bias Estimation

        (a) Segment Level Bias Estimation

           i. CAMS Wheat Estimation Bias
           ii. Ratioed Wheat Estimation Bias

              1. Small Grains Estimation Error
              2. Wheat/Small Grains Ratioing Error

        (b) Country Level — Group III Ratio Error

- Second-Order Error Source Investigations. These investigations examine in further detail the error components identified and/or quantified in the first-order error source evaluations. The second-order error source evaluations are directed toward the investigations of problem areas that have been identified during Phase I and Phase II and toward the examination of error effects that are associated with the operational implementation of new procedures and equipment (e.g., Procedure 1 and the IMAGE 100) in the LACIE Phase III analytical process. A detailed decription

of the second-order source investigations is presented in section 6.2.5.

d. Conduct special investigations of LACIE estimates over selected areas of the U.S.S.R. and Canada. These investigations include:

   - Evaluation of the CAMS proportion estimates for the 30 Canadian test sites and 10 ITS's by comparing them with ground-observed proportions for the same areas. The ground observations are arranged by the Canadian Centre for Remote Sensing.

The specific tasks for implementing the above activities are identified as tasks 10 through 39 in figure 4.1. Specific descriptions of these tasks are provided in section 6.2.

4.3 REPORTING

Reporting for Phase III AA consists of the following three types of reports.

a. Special AA management briefings

b. AA monthly quick-look reports

c. AA interim and final reports

4.3.1 SPECIAL AA MANAGEMENT BRIEFINGS

AA personnel will provide LACIE management with special briefings and presentations on the status of AA data acquisitions and special problems during the 1976-77 winter- and spring-wheat growing seasons. These briefings provide timely responses to management requests for information about LACIE accuracy throughout the growing season.

4.3.2 AA MONTHLY QUICK-LOOK REPORTS

These reports contain an evaluation by AA of the LACIE estimates reported in the Crop Assessment Subsystem (CAS) Monthly Reports (CMR's), the CAS Unscheduled Reports (CUR's), and the CAS Annual Report (CAR). They are released 1 week following the release of a CMR or a CUR if the corresponding SRS or FAS estimates are available. Otherwise, they are released 1 week after the release of the SRS or FAS estimates.

4.3.3 AA INTERIM AND FINAL REPORTS

The interim reports are released in May, August, November, and February; and the final report will be released in April. These reports describe all the results obtained by AA up to the time that each report is written.

The basic reporting formats and the suggested content of these reports are provided in the detailed task descriptions presented in section 6.3 of this plan.

5. SCHEDULE AND RESOURCE REQUIREMENTS

The schedule and resource requirements for implementing LACIE Phase III AA are presented in the following sections.

5.1 SCHEDULE

The Phase III AA schedule is presented in figure 5-1.

5.2 RESOURCE REQUIREMENTS

The resource requirements for LACIE Phase III AA are shown in table 5-1, which summarizes the manpower and computer requirements associated with specific AA tasks or task groups.

TABLE 5-1.– PHASE III ACCURACY ASSESSMENT RESOURCE ESTIMATES

ACCURACY ASSESSMENT RESOURCE ESTIMATES			1976												1977												1978						TASK MAN MONTH			COMPUTER HOURS			
			J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	SF3	SF4	SF12	1108/1110	PDP 1145	IMAGE 100	BENDIX 100
CONTROL MILESTONES		PHASE II REPORTS											3 ▽		4 ▽				F ▽	▽																			
		PHASE III REPORTS																	1 ▽			2 ▽			3 ▽			4 ▽		F ▽									
		TYPE 1 REPORTS																													1 ▽								
311 PHASE III AA PLANNING	SF3	PROJECT INTEGRATION/WRITING										04	04	04	04	04	04	04	04	04	04												4.0						
		TECHNICAL PUBLICATIONS										1	.1	.1	.1	1	1	.1	1	1	1												1.0						
311 AA MONTHLY ANALYSES AND QUICK LOOK REPORTING	SF3	ANALYSIS SESSIONS AND WRITING														.5	5	5	5	.5	5	.5	5	5									45						
		TECHNICAL PUBLICATIONS														1	1	.1	1	1	1	1	.1	1									9						
31.3 AA DATA ACQUISITION AND EVALUATION	SF4	DAPTS DATA MANAGEMENT										10	10	10	10	10	10	10	10	10	10	5	5											11.0					
		CAMS/CAS DATA RETRIEVAL																																					
	SF12	PHOTOGRAPHY – CARTOGRAPHY																		28	28	28	28	28	28	28	28	28	28						28				1806
		PROGRAMMING AND DOCUMENTATION																	30	30	1.05	.25	25	25	.25	25	25								8.55		180 524		
		PROCESSING																																			344		
	SF3	PROGRAMMING AND RESEARCH AND DEVELOPMENT										10	10	10	10	10	10	10	10	10	10	10	.10	10	10	1,0	10	10					17.0				30	40	
		DATA PREPARATION													5	5	5	5	.5	10	20	20	20	2.0	20	20	20	20	2.0	20			23.5						
		PROJECT INTEGRATION											6	6	6	6	6	6	6	6	6	6	6	10	10	10	10	10	1,0	10	10	10	16.2						
		ANALYSES – SPECIAL EARLY SEASON										10	10	20	20	2	.3	4	5	4	4	4	4	5	5	.5	.5	4	4				6.0						
		ANALYSES – LACIE–SRS COMP																															5.8						
		ANALYSES – 90 90 AND FIRST ORDER														6	.8	14	15	1.2	1.2	12	1.2	15	15	1,5	15	1.2	1.2	10	10	.3	19.9			25	20		
		ANALYSES – SECOND ORDER – SEGMENT LEVEL														4	6	9	10	8	8	8	8	10	10	10	10	8	8	7	.7	.2	13.3			25	30		
		ANALYSES – SECOND ORDER – PIXEL LEVEL														8	12	18	20	16	16	16	16	20	20	2.0	2.0	16	1.6	1.3	1.3	5	26.5				150		
3.1.4 AND 3.1.5 AA REPORTING	SF3	WRITING/PRESENTATIONS																		10	10	10	1.0	10	10	1.0	10	10	10	10	10	10	13.0						
		TECHNICAL PUBLICATIONS																		1.3	13	13	13	13	1.3	13	1.3	1.3	1.3	13	1.3	13	16.9						
PHASE III TOTALS																																	162.5		35.55				
PHASE II ANALYSES/WRITING/TECHNICAL PUBLICATIONS												79	79	69	64	58	48	33	28	10													46.8						
TY AA PLANNING	SF3	PROJECT INTEGRATION																				4	.4	20	20	20	20	20	20	4	4	4	14.0						
																						1	.1	1	1	.1	.1	.1	1	1	1	1	1.1						
TY PROCESSING AND AND ANALYSIS ACTIVITIES	SF3	TECHNICAL PUBLICATIONS																				5	5	10	10	10	10	10	10	10	10	10		10.0					
		ANALYSES/EVALUATIONS/REPORTS																						20	26	26	26	36	46	72	92	112	45.6						

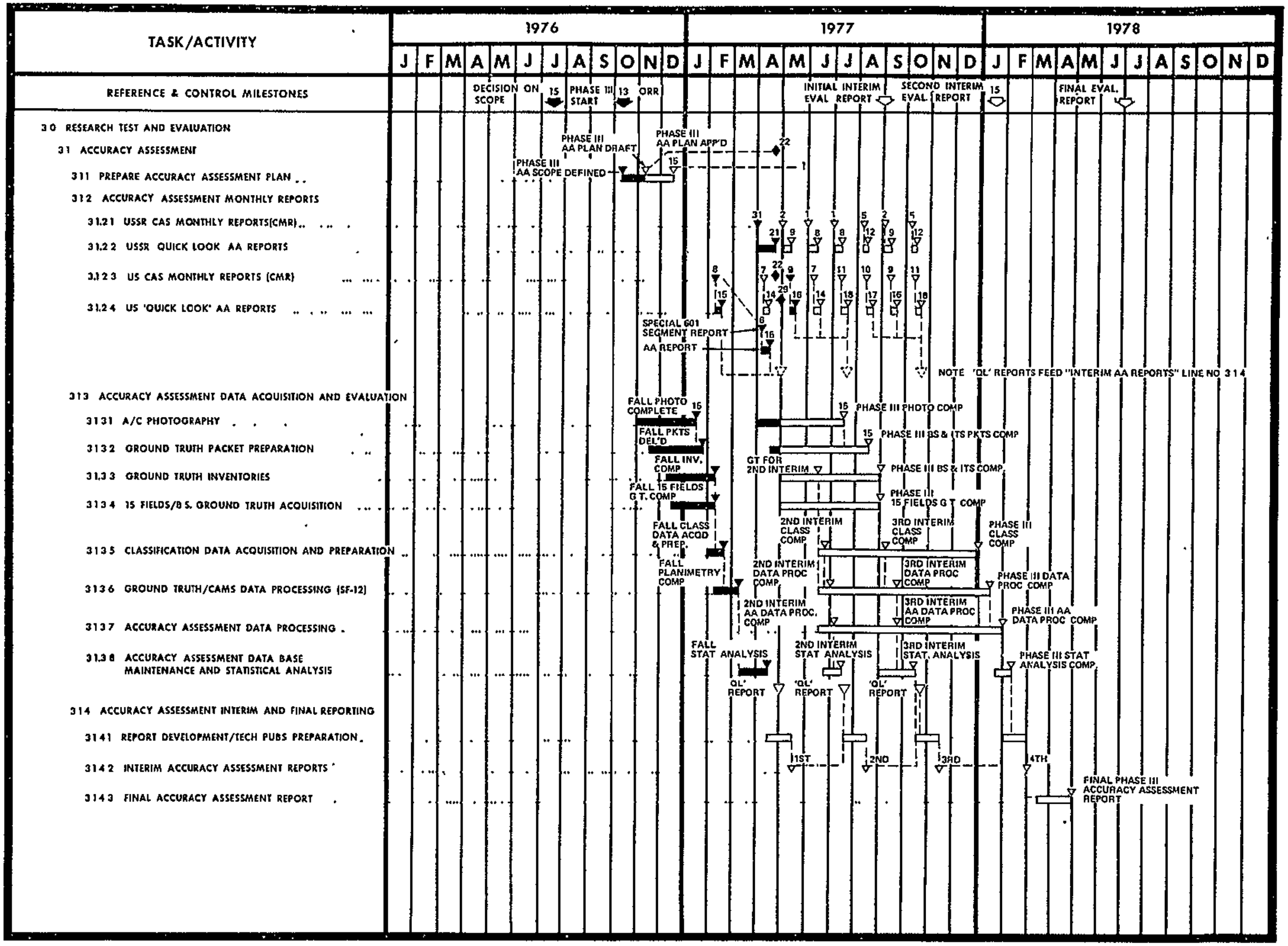


Figure 5-1.— Phase III AA schedule.

ORIGINAL PAGE IS OF POOR QUALITY

PRECEDING PAGE BLANK NOT FILMED

6. TASK DESCRIPTIONS

Detailed descriptions of the tasks that comprise AA for Phase III are included in the following subsections.

6.1 PLANNING AND DATA ACQUISITION

One of the three major elements of Phase III AA is the planning and coordination of analytical and data acquisition activities of the program. This involves (1) the determination of the basic data required to support the program, (2) the coordination and scheduling of data acquisition activities, and (3) the monitoring of data quality to assure that the data acquired are of satisfactory technical quality.

The AA program depends upon the LACIE functional elements [CAMS, CAS, Yield Estimation Subsystem (YES), and Data Acquisition, Preprocessing, and Transmission Subsystem (DAPTS)] to provide a majority of the data necessary for AA evaluations. Specific task descriptions are not included in the AA plan for these LACIE operations. Only those AA data acquisition tasks where direct involvement by the AA Team or by support personnel is required are described in the following paragraphs.

6.1.1 PLAN DEVELOPMENT

The initial activities of Phase III AA are directed to the development of an AA plan. This involves definition of the program scope, the data/resource requirements, the schedules, and the task descriptions of the activities planned for Phase III. The proposed activities of the Phase III AA program are documented in this Phase III AA plan.

6.1.2 DATA REQUIREMENTS

Required inputs from the LACIE functional elements have been identified through the planning for AA investigations. The basic

inputs are defined in the following subsections. As special problems are identified, these requirements may be expanded to satisfy newly defined AA needs.

6.1.2.1 Classification and Mensuration Subsystem (CAMS)

CAMS personnel provide the following information to the AA program for each 1976-77 analysis performed by CAMS and for which results are passed to CAS. This information is also supplied for each classification of each ITS.

a. Segment number.

b. Day of the year (DOY), biostage, and data quality indication for all acquisitions used by an analyst prior to classification.

c. Dot labels and designated other/designated unidentifiable (DO/DU) field definitions prior to classification.

d. DOY, biostage, and data quality indication for any acquisition used by an analyst after classification but prior to evaluation.

e. Dot labels and DO/DU field definitions after classification, if different from (c).

f. Unconditional cluster labels for each pixel as applicable [to be supplied via the Information Storage, Retrieval, and Reformatting Subsystem (ISRRS)].

g. Classifier label applied to each pixel (to be supplied via ISRRS). The bin value of wheat and other categories must be kept constant (e.g., wheat = 239, nonwheat = 143).

h. Prior probabilities used.

i. Classification run number.

j. DOY of the acquisitions used in classification.

k. CAMS evaluation code.

l. Date of analysis.

m. Bias correction applied to classification result.

n. Identification of dots used to label clusters.

6.1.2.2 Crop Assessment Subsystem (CAS)

The following data and information inputs are required from CAS during Phase III AA:

a. U.S. and U.S.S.R. CMR's and any related CAS unscheduled reports (CUR's), which include LACIE Phase III estimates of acreage, yield, and production as they are developed throughout the growing season.

b. The standard statistics for area, yield, and production for LACIE Phase III U.S., U.S.S.R., and Canadian estimates, which include the standard deviation, the CV, the 90-percent confidence limits, and the probability of less than 10 percent relative error.

c. A list of segments used in the aggregation, their estimated percentage of wheat, percentage of small grains, and acquisition date.

d. Wheat/small grains ratios used by CAS in aggregations.

e. Publication of the CAS data base at the end of the crop year (e.g., PC, GPC, etc.)

6.1.2.3 Yield Estimation Subsystem (YES)

The following data and information inputs are required from the YES during the Phase III AA:

a. Adjustable crop calendar (ACC) data used by CAMS to support analyst-interpreter classification activity over the blind sites, 24 U.S. ITS's, and 10 Canadian ITS's.

b. Results from the Phase III operational yield model for each zone in the USGP, for each zone in the U.S.S.R. where yield models exist, and for the years 1964-76 in a stepwise fashion. (This data requirement is similar to that for the Phase I and

Phase II yield feasibility studies.) These results are required to support the first-order yield investigations.

6.1.2.4 Data Acquisition, Preprocessing, and Transmission Subsystem (DAPTS)

The following data and information inputs are required from the DAPTS during Phase III AA.

a. The following ground-observation data for all U.S. blind sites:

- Aircraft photography (the basis for development of field overlays that are used in documenting inventories and interpreting signatures; see section 6.1.4.2 for specific requirements).
- Completed fall early season winter wheat and at-harvest spring and winter wheat inventories of the blind sites conducted by USDA/ASCS personnel using the instructions and data recording forms presented in appendix D.
- Eighteen-day observations of crop height and ground cover over 15 wheat fields in each blind site.

b. The following ground-observation data (to be collected as a part of LACIE operations) for all U.S. ITS's:

- Aircraft photography (the basis for development of field overlays that are used in documenting inventories).
- Completed fall wall-to-wall and spring inventories of the ITS's conducted by USDA/ASCS personnel utilizing the data forms presented in appendix E.
- Completed 18-day periodic observations of approximately 50 fields from the ITS's throughout the wheat growing season until harvest, to be taken within 3 days of Landsat overflights. (The USDA/ASCS personnel make and record these observations on the forms presented in appendix E.)

c. Ground observation of wheat growth-stage data (after each growth stage) acquired by USDA/ASCS personnel over U.S. ITS's.

d. Copies of mean historical crop calendars based on the last 15 years, if available, for each acreage stratum for the USGP states and Canada and copies of the ACC's, as available every 2 weeks by CRD.

e. Agriculture reports:

- USDA/SRS reports containing current information on wheat acreage, yield, and production for the United States at the state level. These shall be made available to AA personnel on the day of release.

- USDA/FAS reports containing current year information on wheat acreage, yield, and production for Canadian and U.S.S.R. wheat growing areas. These shall be made available to AA personnel on the day of release.

f. Historical agricultural statistics:

- USDA/SRS data on wheat acreage, yield, and production in the United States for 1970-75.

- USDA agricultural census data for 1969 and 1974.

g. Other required data sets as specified by AA personnel to satisfy special investigations that may be requested by LACIE project management. AA personnel will specify any such requirements to DAPTS as soon as possible after the data requirements are identified.

6.1.3 SITE SELECTION

For Phase III, the AA team will select 143 winter wheat and 69 spring wheat blind sites in the United States and 30 test sites in Canada.

6.1.3.1 U.S. Blind Sites

The U.S. blind sites were selected during the fall of 1976 so that aircraft photography could be obtained earlier than was the case in Phase II. The location and identity of all U.S. blind sites will remain unknown to the CAMS data analysts so that these sites can be processed as regular segments. The 143 winter wheat blind sites were selected by a random draw stratified by CRD from five states in the USSGP (Colorado, Kansas, Nebraska, Oklahoma, and Texas) and from two mixed wheat states (South Dakota and Montana). The 69 spring wheat blind sites were selected from the two mixed wheat states (South Dakota and Montana) and from the two spring wheat states (North Dakota and Minnesota). The distribution of blind sites by state is included in appendix B (see table B-3).

6.1.3.2 Canadian Spring Wheat Test Sites

Thirty spring wheat test sites will be selected from the Province of Saskatchewan. These sites are similar to blind sites except that their identity is known to the analysts.

6.1.4 PREPARATION OF BLIND SITE FIELD OVERLAYS

Field overlays will be prepared from aircraft photographs of the blind sites and will be used to record the land-use information obtained by observation on the ground. The following items are required to prepare the blind site field overlays.

6.1.4.1 Aircraft Maps

After selection of the blind sites, Landsat imagery is used to determine the true position of each site. Analyst-interpreters determine these positions using production film converter (PFC) products, record the latitude and longitude to the nearest

0.1 minute, and plot the position of the segment on a 1:24 000-scale or 1:12 500-scale map. These maps are then used by air-crews in acquiring the aerial imagery.

6.1.4.2 Aircraft Photography

Aerial photography is collected using color infrared film. If possible, this photography is obtained from a high altitude so that a single photograph covers the entire site. If this is not possible, the flight is made at an altitude of 6000 to 7200 meters (20 000 to 24 000 feet) and two flight lines are flown for each site with a 20-percent sidelap. Four frames are collected for each flight line with a 30-percent forward overlap. All imagery must be collected no later than 4 weeks prior to ground-truth collection. Predesignated flight lines are established for each blind site.

After aerial imagery is acquired for the blind sites, each frame is checked to verify that the site was covered and that the imagery is of sufficient quality to be used by the USDA/ASCS personnel in collecting ground-truth data.

6.1.4.3 Field Overlays and Field Segment Kits

If the imagery is of satisfactory quality to be used by USDA/ASCS personnel, transparent overlays are prepared. The overlays are then placed in field segment kits that are forwarded to USDA/ASCS personnel in the appropriate county for use in acquiring ground-truth data. These kits include:

- A color infrared 2X or 4X print of the segment with boundaries on the field overlays. (The 4X enlargement is used for high-altitude photography.)
- A topographical map of scale 1:250 000 showing the sample segment location and boundaries.

- Crop identification key with standard annotation for documenting land use.
- Survey manual with a brief definition of field procedures developed at JSC providing guidelines to USDA/ASCS personnel for recording ground observations of the LACIE blind sites (appendix D).

6.1.5 BLIND SITE FIELD DATA ACQUISITION

USDA/ASCS personnel provide complete inventory data based on ground observations. The data for each field are annotated on the overlay according to the standard crop symbols identified in the crop keys provided in the JSC instructions to USDA/ASCS for making LACIE segment inventories (appendix D). These inventory packages are to be completed by USDA/ASCS personnel and forwarded to JSC to be logged and tracked by DAPTS.

All blind sites in the USSGP that have an early season planted inventory will have 15 wheat fields chosen and annotated on the overlay by USDA/ASCS personnel: 5 below average stands, 5 average stands, and 5 above average stands. The USDA/ASCS personnel will identify the plant height and ground cover of each of these fields at this time. Beginning on April 6, the USDA/ASCS personnel will begin to revisit these 15 fields in concert with the Landsat overpasses so that classification performance can be related to wheat field stands. Also beginning April 6, similar observations will commence over all blind sites which have been planted.

As discussed in appendix H, software is being developed which will be run in the background mode on the PDP 11/45 and will determine, for each of these fields, the amount classified as wheat and the amount classified as other for each classification.

Prior to the time that this software is ready, manual interpretation is necessary. To do this, AA personnel will ask CAMS to pull the packets; AA will pull the 18-day field observations and the ground-truth overlay. The latest imagery used in classification and the classification map from the packet will be projected onto the ground truth overlay, and the percentage of each of the 15 fields classified as wheat will be estimated and recorded.

Fields 1 through 15, recorded by USDA/ASCS personnel, are assigned numbers ranging from 300 through 314 on the ground-truth overlay. For those segments with training dots defined on the appropriate PFC, each dot or small-field group of four dots will be verified in terms of the label given by the analyst versus the ground-truth label. Likewise, areas of other crops which are classified as wheat are estimated by AA personnel. When the automated system becomes operational, these manual interpretations will be used for order-of-magnitude verification tests.

6.1.6 PROCESSING BLIND SITE GROUND-TRUTH DATA TO SUPPORT SEGMENT-LEVEL AA INVESTIGATIONS

The early season blind site ground-observation data will be processed according to the procedures used in Phase II. These procedures require the LACIE cartographic technician to plot the LACIE segment boundary on a product 1, 2X photograph. Using the area mode feature of the H. Dell Foster digitizer, the technician planimeters or measures the segment area in thousandths of a square inch on the photograph. Next, the proportions of wheat, small grains, abandoned wheat, and abandoned small grains (in each case separated into spring and winter) are determined by planimetering the area for each of these classes and dividing by the total area in the segment. These proportions are used in various investigations described in section 6.2.

The late season blind site ground-observation data will be processed in two ways to determine proportions. The first method is designed to obtain quick estimates of the proportions and the second (slower) method is designed to obtain the exact proportions.

The first method involves placing a grid containing 400 dots over the ground-truth annotated aircraft imagery and determining the class of each dot. The proportions in the image are then assumed to be the same as the proportions in the dot samples. It is expected that the standard error in this procedure will be ±2.4 percent. Proportions will be determined for wheat and small grains, and these classes will be further broken down into the categories of spring, winter, harvested, and abandoned. The following list gives the schedule for completing these estimates:

a. North Dakota (27) — August 25

b. South Dakota (19) — August 31

c. Oklahoma (20 — September 7

d. Remaining USNGP (69) — September 19

e. Colorado (13) — September 21

f. Remaining USSGP (41) — September 26

g. Canadian Sites (30) — October 1

Here the number in parentheses is the number of segments.

The second method for determining ground-truth proportions involves using the Bendix 100 to determine the vertices of each field in the segment. These will be stored on a magnetic tape along with the identification of the ground-truth class for each field. The proportions will then be calculated by the computer routine SPECTL, which is part of the pixel-level processing system described in appendix H.

6.1.7 FIELD DATA ACQUISITION IN INTENSIVE TEST SITES

The field data acquisition from 24 U.S. and 10 Canadian ITS's is an integral part of LACIE operations. These sites are located prior to Phase III operations, and their identities and locations are available to all LACIE personnel (see appendix B). Field data acquired from these sites by USDA/ASCS personnel include the following:

a. Aerial photography (once yearly)

b. Field maps annotated by USDA/ASCS personnel

c. Inventories of all fields (Figure E-1 in appendix E provides an example of the ground-truth data reporting forms.):

- After fall planting for winter wheat areas
- "At harvest" for spring and winter wheat areas

d. Periodic 18-day observations of a subsample (approximately 50 fields) of each ITS coincident with each Landsat overpass (Figure E-2 in appendix E gives an example of the ground-truth periodic observation form used for recording these 18-day periodic observations.)

These data are forwarded to JSC to be processed, logged by DAPTS, and stored in the LACIE Physical Data Library (LPDL) where this information is then made available to AA personnel.

6.1.8 PROCESSING ITS GROUND-TRUTH DATA TO SUPPORT SEGMENT-LEVEL INVESTIGATIONS

Measurement of wheat and small grain proportions (both spring and winter) in the ITS will be done by using the Phase II procedure of adding field acreages from the inventory list.

6.2 ANALYSIS AND EVALUATION

The analysis and evaluation tasks described in this section are those presently planned to accomplish the Phase III AA objectives described in section 1.1.

6.2.1 DATA COLLECTION AND PREPARATION

This task involves collecting the data required by AA from LACIE operations and other sources. It is listed as a separate task because it involves considerable effort on the part of AA. The various data sets required are shown in figure 4-1.

6.2.2 COMPARISONS OF LACIE ESTIMATES WITH REFERENCE STANDARDS AS A METHOD OF ASSESSING THEIR ACCURACY AND RELIABILITY DURING EARLY SEASON AND THROUGHOUT THE GROWING SEASON

A prime concern of the LACIE AA program is to monitor and evaluate estimates made during early season and at regular intervals throughout the growing season. These evaluations are made through comparisons with the reference standard (USDA/SRS estimates).

The statistic used for making these comparisons between the LACIE estimates of wheat production, area, and yield and the corresponding reference estimates is the relative difference (RD) defined by:

$$RD = \frac{\text{LACIE} - \text{STANDARD}}{\text{LACIE}} \times 100\%$$

where LACIE stands for the LACIE estimate of wheat production, area, or yield and STANDARD represents the corresponding reference standard estimate. This definition expresses the difference between the two estimates as a percentage of the LACIE estimate. In the United States these comparisons are made for each state in the USGP and for the various regions discussed in section 4.2. In the U.S.S.R., they will be made at the country level and possibly for certain regions.

Significance tests of no difference are made at the region or country level for the LACIE production, area, and yield estimates for spring wheat, winter wheat, and total wheat. In order to make a significance test, the LACIE estimate (of wheat production, area, or yield) is assumed to be approximately normally distributed with unknown mean μ and variance σ^2_{LACIE}. A test of the hypothesis

$$H_o: \quad \mu = \text{STANDARD}$$

versus the alternative hypothesis

$$H_A: \quad \mu \neq \text{STANDARD}$$

is then made using this assumption. The test statistic is given by

$$Z = \frac{\text{LACIE} - \text{STANDARD}}{\hat{\sigma}_{\text{LACIE}}}$$

which, under the null hypothesis, is approximately normally distributed with mean zero and variance one. The null hypothesis is rejected in favor of the alternative at the α-level of significance if

$$|Z| > z_{\alpha/2}$$

where $z_{\alpha/2}$ is the $\left(1 - \frac{\alpha}{2}\right)$ critical point of the standard normal distribution. For $\alpha = 0.10$, $z_{\alpha/2} = 1.645$, and if $|Z| > 1.645$, it is concluded that the mean of the LACIE estimator is significantly different from the reference standard estimate.

These comparisons are designed to detect any abnormal divergences between the estimates and reference standards and thus to identify for further investigations of potential LACIE problem areas that might be associated with the divergences.

The results of these evaluations are reported in the Phase III AA monthly quick-look reports and in the interim and final AA reports which are described in sections 6.3.2 and 6.3.3, respectively.

6.2.3 DETERMINING IF LACIE PRODUCTION ESTIMATES MEET THE 90-90 CRITERION

Let $\hat{P}$ be the LACIE estimate of wheat production for the region or country, and let P be the true wheat production of the same region or country. The accuracy goal of the LACIE is a 90-90 at-harvest criterion for wheat production, which is given by the following probability statement.

$$\Pr[|\hat{P} - P| \leq 0.1P] \geq 0.90 \tag{1}$$

This states that the accuracy goal is for the LACIE at-harvest estimate of wheat production to be within 10 percent of the true wheat production with a probability of at least 0.9.

It is assumed that the LACIE estimate, $\hat{P}$, is normally distributed with mean P + B and variance $\sigma^2_{\hat{P}}$, where B is the bias given by

$$B = E(\hat{P}) - P$$

Under this assumption, equation (1) may be written as

$$\Pr\left[\frac{-0.1 - 0.9\frac{B}{P + B}}{CV(\hat{P})} \leq Z \leq \frac{0.1 - 1.1\frac{B}{P + B}}{CV(\hat{P})}\right] \geq 0.90 \tag{2}$$

where $Z = \frac{P - (P + B)}{\sigma\hat{P}}$ follows the standard normal distribution, N(0,1), and $CV(\hat{P})$ is the coefficient of variation of $\hat{P}$ defined by

$$CV(\hat{P}) = \frac{\sigma_{\hat{P}}}{E(\hat{P})} = \frac{\sigma_{\hat{P}}}{P + B} \tag{3}$$

The term $\frac{B}{P + B}$ is called the relative bias of $\hat{P}$ and is given by

$$\frac{B}{P + B} = \frac{E(\hat{P}) - P}{E(\hat{P})} \tag{4}$$

It follows that the accuracy goal of LACIE is attained if

$$\Phi\left[\frac{0.1 - 1.1RB(\hat{P})}{CV(\hat{P})}\right] - \Phi\left[\frac{-0.1 - 0.9RB(\hat{P})}{CV(\hat{P})}\right] \geq 0.90 \qquad (5)$$

where Φ represents the cumulative standard normal distribution. The enclosed region of figure 6-1 indicates combinations of $CV(\hat{P})$ and relative bias for which equation (5) is satisfied.

6.2.3.1 Variance and Bias Estimation for the Wheat Production Estimate

To apply the evaluation technique described in the previous section, knowledge of the variance, $\sigma^2_{\hat{P}}$, and bias, B, of the LACIE wheat production estimate for a country or region is required. Since values of these parameters are unknown in LACIE, their estimates have to be obtained. The estimation of the production variance at different aggregated levels is described in detail in reference 1 and will not be discussed here. An estimate of the bias can be obtained by noting the difference between the LACIE production estimate and the corresponding USDA estimate of production. But this is possible only in the United States using USDA/SRS estimates. For foreign countries, USDA/FAS makes periodic forecasts which are generally for total grains production and are assessed using ad hoc methods. Although USDA/FAS estimates may be utilized to note any major problem, they cannot be used for a quantitative assessment of bias in a LACIE estimate.

6.2.3.2 At-Harvest 90-90 Criterion Evaluation

Given the estimator for $\sigma^2_{\hat{P}}$, an estimate of $CV(\hat{P})$ is

$$\hat{CV}(\hat{P}) = \sigma^2_{\hat{P}}/P$$

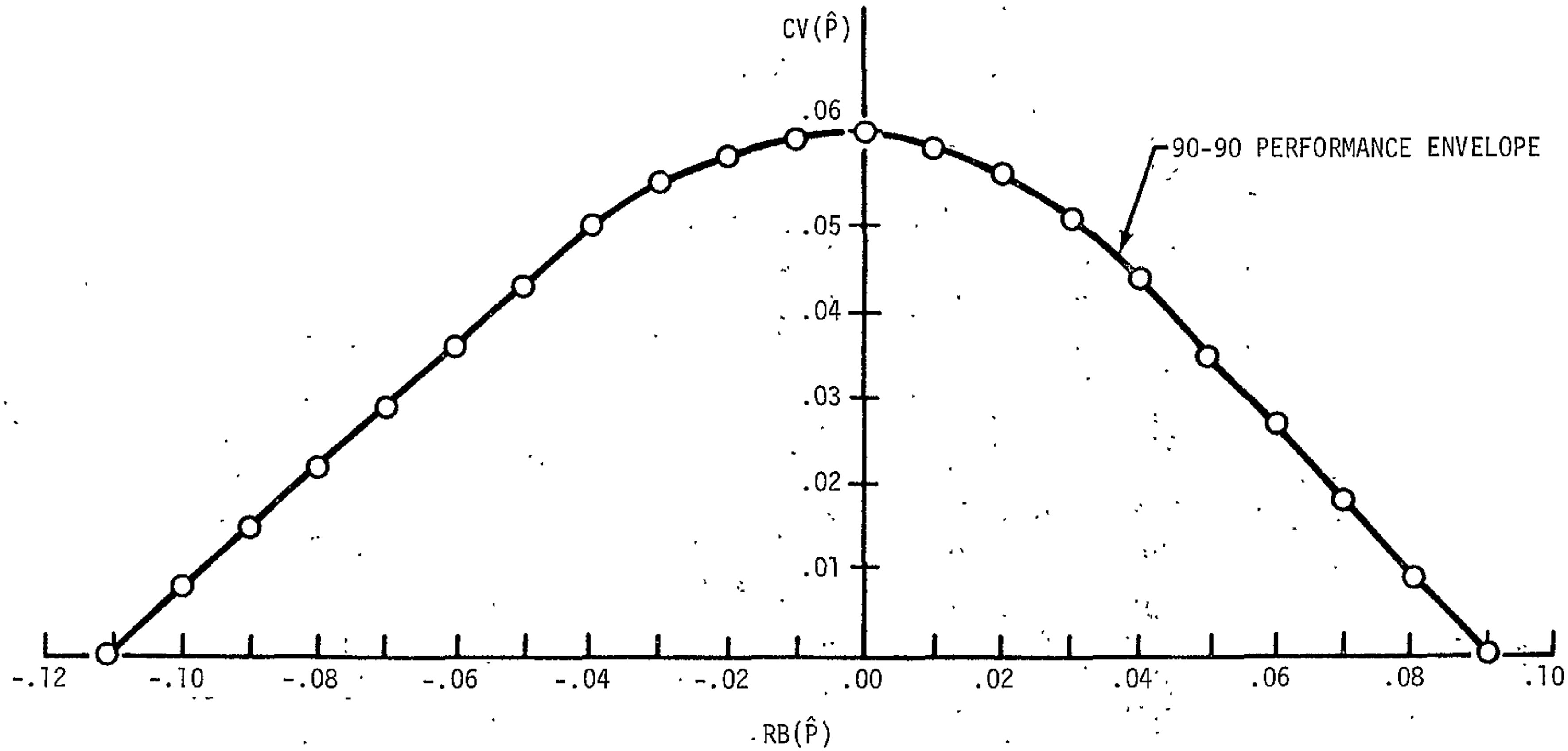


Figure 6-1.— Relative bias versus coefficient of variation (CV) of production.

where the computation of $\sigma^2_{\hat{P}}$ is described in detail in reference 1 and is provided by the CAS aggregation software. An estimator of $RB(\hat{P})$ is

$$\hat{RB}(\hat{P}) = \hat{B}/\hat{P}$$

where $\hat{B}$ is the difference between the LACIE production estimate and the corresponding USDA/SRS estimate.

The observed value of the left side of equation (5), above, with $CV(\hat{P})$ and $RB(\hat{P})$ replaced by their estimates, $\hat{CV}(\hat{P})$ and $\hat{RB}(\hat{P})$, respectively, is subject to certain variability, which is intractable due to problems in obtaining a joint distribution of $\hat{CV}(\hat{P})$ and $\hat{RB}(\hat{P})$. However, if $\hat{CV}(\hat{P})$ is greater than 0.061, there is an indication that the LACIE estimator does not satisfy the 90-90 criterion even if $\hat{P}$ is unbiased. Since $\hat{CV}(\hat{P})$ has been found to be very stable at the country level (USGP level in the case of the United States) and less than 0.061, $\hat{CV}(\hat{P})$ is treated as the parameter $CV(\hat{P})$, and equation (5) can be solved to determine the tolerable values of $RB(\hat{P})$ that would meet the 90-90 accuracy goal. That is, given $CV(\hat{P})$, there exist real numbers $R_o (R_o > 0)$ and $R_1 (R_1 > 0)$ so that equation (5) is satisfied for

$$R_o \leq RB(\hat{P}) \leq R_1$$

or, equivalently,

$$B_o \leq B \leq B_1$$

where $B_o = R_o P/(1 - R_o)$ and $B_1 = R_1 P/(1 - R_1)$, P being the true production.

Suppose next that the LACIE production estimator is a 90-90 estimator; i.e., suppose $CV(\hat{P}) \equiv \hat{CV}(\hat{P}) \leq 0.061$ and $RB(\hat{P}) \in [R_o, R_1]$.

Under this assumption, calculate the significance level

$$SL = \max_{RB(\hat{P}) \in [R_0, R_1]} \{\min[\text{Prob}[\hat{RB}(\hat{P}) < rb(\hat{P})], \text{Prob}[\hat{RB}(\hat{P}) > rb(\hat{P})]]\}$$

$$= \max_{B \in [B_0, B_1]} \{\min[\text{Prob}[\hat{B} < b], \text{Prob}[\hat{B} > b]]\} \qquad (6)$$

where $rb(\hat{P})$ and b are the observed $\hat{RB}(\hat{P})$ and $\hat{B}$, respectively, and the two probabilities are computed assuming $\hat{B}$ is normally distributed with mean B and variance $\sigma^2_{\hat{P}}$. Since $\hat{P}$ is assumed normally distributed with mean (P + B) and variance $\sigma^2_{\hat{P}}$,

$$\hat{B} = \hat{P} - P_{SRS}$$

is also normally distributed with mean B and variance $\sigma^2_{\hat{P}}$, provided $P = P_{SRS}$. The estimated significance level, SL, is an estimate of the probability of encountering the observed difference, given that the LACIE production estimator is a 90-90 estimator. If SL is small, say less than 0.10, it is concluded that the 90-90 accuracy goal was not attained due to a consistent bias that is larger than the tolerable amount. If SL is larger than 0.10, it is *not* immediately inferred that the LACIE production estimator is a 90-90 estimator. This is due to the fact that only one observation has been obtained to estimate B. However, results obtained from blind site analyses and other AA tasks are then considered for further assessment of whether or not the 90-90 accuracy goal is achievable.

6.2.3.3 Early Season 90-90 Criterion Evaluation

Although the "official" evaluation of the 90-90 criterion is based on at-harvest estimates, it is of interest to evaluate how well LACIE is performing throughout the season. When data are available for both spring and winter wheat (generally after July), the evaluation is performed in the same way as for the at-harvest

estimate. In order to gauge how well LACIE is performing early in the season when only winter wheat data are available, a method was developed to project the winter wheat results for the 5- or 7-state level to the 9-state total harvestable wheat level.

The 5- or 7-state relative difference between the LACIE and USDA estimates is taken as an estimate of the relative bias. The coefficient of variation, however, is "projected" to the 9-state level by

$$CV' = CV_R \sqrt{\frac{N_R}{N_{US}}} \tag{7}$$

where CV_R is the current month $\hat{CV}(\hat{P})$ for the 5- or 7-state winter wheat production estimate, and N_R and N_{US} are the corresponding numbers of allocated segments for the 5- or 7-state region and the USGP region, respectively. After the relative bias and coefficient of variation have been estimated, inference as to whether the 90-90 criterion has been supported is made using the evaluation procedure discussed in the previous section.

6.2.4 ERROR SOURCE EVALUATIONS — FIRST-ORDER ERROR SOURCE INVESTIGATIONS

A major purpose of AA is to attempt to ascertain the nature and characteristics of the error in LACIE Phase III production estimates. This requires indepth investigations in which the error in LACIE production estimates is quantitatively and/or qualitatively associated with various causative factors. However, the error in production depends on its sources in a complex way; thus, it is unrealistic to assume that the total error component can be written as a sum of uncorrelated random components. Instead, the effects of the major components are evaluated by estimating the reduction in the prediction error of production, which results from eliminating that component of error. These major

TABLE 6-1.— PHASE II CV'S AND RELATIVE DIFFERENCES

Date	Area, state	Phase II CV for production	Phase II relative difference
Feb.	5	11	-4.9
Mar. 25	5	10	-9.9
Apr. 8	5	8	-8.5
May 7	5	8	-1.6
June 8	5	7	11.4
	7	8	1.7
June 29	5	7	12.7
	7	7	4.7
July 9	5	7	-3.7
	7	7	-7.9
Aug. 11	5	7	-4.2
	7	7	-5.6
	9	6	-14.7
Sept. 9	5	7	-6.6
	7	7	-6.6
	9	5	-13.6
Oct. 8	5	7	-6.6
	7	7	-6.5
	9	5	-13.8
Dec. 17 (final)	5	7	-7.2
	7	7	-7.2
	9	5	-12.3

components of production error are called first-order errors. They are graphically depicted in figure 6-2. These first-order errors are the errors contributing to the LACIE production estimate which can be quantitatively estimated from LACIE estimates, reference standard estimates (USDA/SRS), and historical (county census) and blind site data. Methods for evaluating these first-order error components are discussed in sections 6.2.4.1 through 6.2.4.5. Section 6.2.4.6 describes the method for determining the effect of errors in yield, acreage, sampling, and classification on the production estimate.

6.2.4.1 Proportion Estimation Errors

Proportion estimation errors are determined for the CAMS estimates of small grains and ratioed wheat. If CAMS estimates wheat directly (i.e., without using ratio) in Phase III, the proportion estimation errors for these estimates will also be studied.

The proportion error for a given blind site segment is $\hat{X} - X$, where $\hat{X}$ is the CAMS proportion estimate and X is the ground-truth proportion. These errors are studied by plotting them as a function of X for each state in the USGP and for the USGP region. Also, statistics for these regions are calculated as follows.

Let N be the number of segments allocated to a region (state or higher level), and let n be the number of blind sites selected randomly from these N segments. For a region, let $\hat{X}_i$ represent the CAMS estimate of the proportion of wheat in the *ith* segment and let X_i represent the ground-truth proportion of wheat in the *ith* segment, where $i = 1, \cdots, N$. Then the average error μ_D is given by

$$\mu_D = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{X}_i - X_i\right) \qquad (8)$$

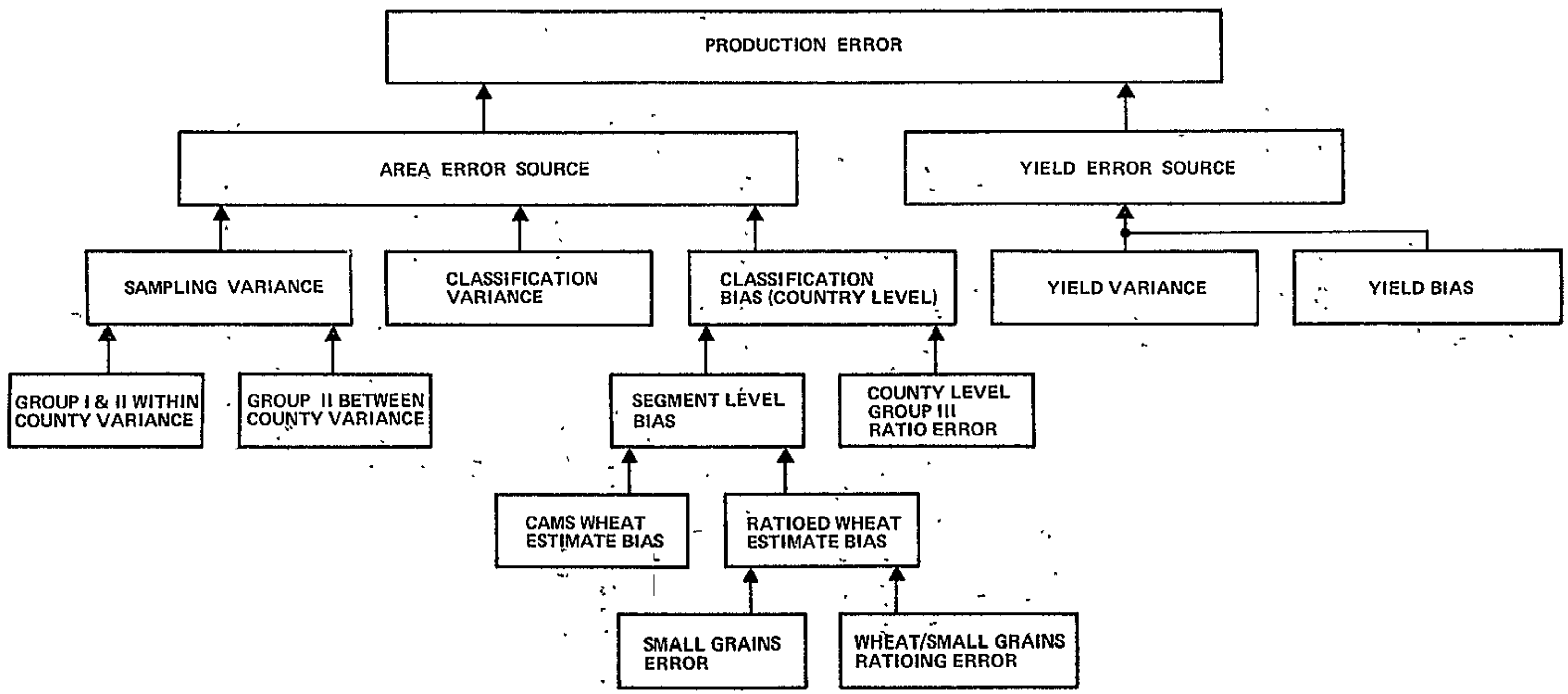


Figure 6-2.— LACIE first-order error components.

The estimate of μ_D is given by

$$\bar{D} = \frac{1}{n} \sum_{i=1}^{n} (\hat{X}_i - X_i) \tag{9}$$

where the summation is taken over the n blind sites.

Letting $D_i = \hat{X}_i - X_i$, we may estimate the variance of $\bar{D}$ by

$$S_{\bar{D}}^2 = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{\sum_{i=1}^{n} (D_i - \bar{D})^2}{n - 1} \tag{10}$$

Lower and upper confidence limits for the population average difference $\bar{D}_p$ are given by

$$\mu_{D_L} = \bar{D} - t_{1-\alpha/2} S_{\bar{D}} \; ; \quad \mu_{D_U} = \bar{D} + t_{1-\alpha/2} S_{\bar{D}} \tag{11}$$

where $t_{1-\alpha/2}$ is the value of $1 - \alpha/2$ percentage point, from the Student's t distribution with (n - 1) degrees of freedom, corresponding to the desired confidence level of $1 - \alpha$.

The null hypothesis $\mu_D = 0$ (i.e., no bias) is rejected at the α-level of significance if $|\bar{D}/S_{\bar{D}}| > t_{1-\alpha/2}$, or equivalently, if the confidence interval given by equation (7) does not contain zero.

The quantities $\bar{D}$, $S_{\bar{D}}$, μ_{D_L}, and μ_{D_U} are tabulated at the state and the USGP levels, and the test is used to determine whether or not there is bias at these levels.

6.2.4.2 Contribution of the Classification and Ratioing Errors to the Ratioed Wheat Proportion Estimation Errors at the Segment Level

This section describes the method used to study the contribution of the classification and ratioing errors to the ratioed wheat proportion errors, which consist of the proportion bias and the proportion mean-square error (MSE).

Let $\hat{r}_i$ and $\hat{x}_i$, $i = 1,2,\cdots,n$, be the estimates of r_i and x_i respectively, for the *ith* blind site, where r_i is the ground-observed ratio of wheat to small grains proportion; x_i is the ground-observed small grains proportion; and n is the number of blind sites. In this section, $\hat{r}_i$ is the forecast ratio of wheat to small grains proportions, and $\hat{x}_i$ is the CAMS estimate of the small grains proportion.

The bias (B) and the MSE of the wheat proportion estimate obtained after ratioing may be estimated by

$$\hat{B} = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{r}_i \hat{x}_i - r_i x_i\right) \tag{12}$$

and

$$\hat{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{r}_i \hat{x}_i - r_i x_i\right)^2 \tag{13}$$

It is clear that these errors are both caused by two factors: the CAMS classification of small grains and the estimated ratio of wheat to small grains. The contribution of a particular error factor may be assessed by the reduction in the bias or MSE which would be achieved if that error factor were omitted. Specifically, the following formulas are used in this study.

a. Bias estimate with no ratioing error:

$$\hat{B}' = \frac{1}{n} \sum_{i=1}^{n} \left(r_i \hat{x}_i - r_i x_i\right) \tag{14}$$

b. Bias estimate with no classification error:

$$\hat{B}'' = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{r}_i x_i - r_i x_i\right) \tag{15}$$

c. MSE estimate with no ratioing error:

$$\hat{MSE}' = \frac{1}{n} \sum_{i=1}^{n} \left(r_i \hat{x}_i - r_i x_i\right)^2 \tag{16}$$

d. MSE estimate with no classification error:

$$\hat{MSE}'' = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{r}_i x_i - r_i x_i\right)^2 \tag{17}$$

These quantities will be calculated at the state and USGP levels, and a sensitivity analysis will be conducted to measure the effect of classification and ratio error on the bias and MSE for ratioed wheat proportion.

Data required for this investigation are:

a. CAS ratios for every USGP blind site used in the aggregation.

b. CAMS estimates of spring and winter wheat and either a small grains or a spring or winter small grains estimate for every 1977 blind site.

c. Ground-truth proportions for spring wheat, winter wheat, small grains, winter small grains and spring small grains for every 1977 USGP blind site.

6.2.4.3 Classification Bias

The LACIE estimate of wheat acreage for a large area in the United States can be written

$$\hat{A} = \sum_{i=1}^{n} W_i \hat{X}_i \tag{18}$$

where $\hat{A}$ is the estimated wheat acreage of the region; $\hat{X}_i$ is the wheat proportion estimate in the *ith* LACIE segment, n is the number of processed LACIE segments; and the W_i are weights based on historical and cartographic data.[1]

Corresponding to $\hat{A}$ is the true acreage, A, which can be written

$$A = \sum_{i=1}^{n} W_i^* c_i \tag{19}$$

where c_i is the true wheat acreage for the county containing the *ith* segment and W_i^* is the value of the weight which would give perfect Group III estimates of wheat acreage for unsampled counties.

We can now write

$$\begin{aligned} \hat{X}_i &= c_i + (x_i - c_i) + (\hat{X}_i - X_i) \\ &= c_i + \delta_i + \varepsilon_i \end{aligned} \tag{20}$$

where X_i is the true wheat proportion of the *ith* segment, δ_i is the sampling error, and ε_i is the classification error. Since the sampling was performed in an unbiased manner, we assume

---

[1] The precise definition of W_i depends on whether the *ith* segment is used as part of a Group III estimate.

$E(\delta_i) = 0$; however, we do not assume unbiased classification. Instead, let θ be an average segment bias; i.e.,

$$E(\varepsilon_i) = \theta \tag{21}$$

The bias in $\hat{A}$ is defined by $E(\hat{A} - A)$, which is given by

$$\begin{aligned} B = E(\hat{A} - A) &= E\left(\sum_{i=1}^{n} W_i X_i - \sum_{i=1}^{n} W_i^* c_i\right) \\ &= \sum_{i=1}^{n} W_i E(c_i + \delta_i + \varepsilon_i) - \sum_{i=1}^{n} W_i^* c_i \\ &= \sum_{i=1}^{n} (W_i - W_i^*) c_i + \theta \sum_{i=1}^{n} W_i \end{aligned} \tag{22}$$

Note that the first term of equation (22) represents a bias caused by the failure of the Group III ratios to be exact (i.e., $W_i \neq W_i^*$); whereas the second term is the average segment bias multiplied by the sum of the W_i.

At present, only the second term of equation (22) is estimated since good county-level data are not available for estimating the first term.

The second term is estimated by (1) breaking up the large area into strata (not necessarily connected) for which equation (21) holds; i.e., the bias is constant, (2) estimating θ by $\hat{\theta} = \frac{1}{n_k} \sum_{i=1}^{n_k} (\hat{X}_i - X_i)$, the average proportion error on a segment level in the k*th* stratum where n_k is the number of blind sites in the k*th* stratum, and (3) aggregating $\hat{\theta}$ over the strata.

If $\hat{B}$ represents the estimate of bias for the region resulting from classification, a 90-percent confidence interval for B (the true bias) can be constructed by

$$(\hat{B} - 1.645\hat{\sigma}, \hat{B} + 1.645\hat{\sigma})$$

where $\hat{\sigma}$ is an estimate of the standard error of $\hat{B}$.

If we assume $\mathrm{Var}(\varepsilon_i) = \sigma_{ck}^2$ (a constant) within the k*th* stratum, then σ_{ck}^2 can be estimated by $\sum_{i=1}^{n_k} \frac{(\hat{x}_i - x_i - \hat{\theta})^2}{n_k - 1}$ and $\mathrm{Var}(\hat{B})$ can be estimated by $\hat{\mathrm{Var}}(\hat{B}) = \sum_k \hat{\sigma}_{ck}^2 \left(\sum_{i=1}^{n_k} W_{ki} \right)$, where W_{ki} is the weight for the i*th* segment in the k*th* stratum.

6.2.4.4 Estimation of the Within-County Acreage Variances Resulting From Classification and Sampling Errors

In order to estimate the within-county acreage variances resulting from sampling and classification errors, one first obtains three basic regression models; namely, true segment proportion versus historical county proportion, LACIE segment proportion versus ground-truth segment proportion, and LACIE segment proportion versus historical county proportion. Then these regression equations are used to obtain the estimates for $\sigma_s^2 + \sigma_H^2$, σ_c^2, and $\lambda^2\sigma_s^2 + \sigma_c^2$ where $\lambda^2\sigma_s^2$, σ_c^2, and σ_H^2 represent, respectively, the contribution resulting from sampling, the contribution resulting from classification, and the variance of the residuals resulting from the regression of the current county proportion onto the historical county proportion. (Assuming that σ_H^2 is much smaller than σ_s^2, σ_H^2 can be ignored in the calculation.) Finally, the maximum likelihood estimation technique, assuming normality, is used to obtain the optimal estimates for sampling and classification variances. The detailed description of this method is presented in appendix C.

6.2.4.5 Yield Bias and Variance

To support the evaluation of the 90-90 criterion for production, the yield estimates must be tested for bias and the accuracy of the corresponding estimated variances determined. In order to do this, the bias and variance are tested, as first-order error source evaluations, for each zone for all truncations for the U.S.S.R. and the USGP using 10 or more years of independent test data. This is being done as a test and evaluation task. The data to be used are as follows.

a. United States 1965-76

b. U.S.S.R. 1958-72

c. Canada 1967-76

6.2.4.6 Production Bias

The production bias at the state level is given by

$$\begin{aligned} B_{P_i} &= E(\hat{P}_i - P_i) \\ &= E(\hat{P}_i) - P_i \\ &= E(\hat{A}_i\hat{Y}_i) - A_iY_i \end{aligned} \tag{23}$$

where P_i, A_i, and Y_i are the true values of the production, acreage, and yield, respectively, for the *ith* state in question, and $\hat{P}_i$, $\hat{A}_i$, and $\hat{Y}_i$ are the corresponding estimates for these quantities. Assuming A_i and Y_i are independent, one obtains

$$B_{P_i} = E(\hat{A}_i)E(\hat{Y}_i) - A_iY_i \tag{24}$$

If one further assumes that $\hat{Y}_i$ is unbiased, then $E(\hat{Y}_i) = Y_i$, and

$$\begin{aligned} B_{P_i} &= Y_i[E(\hat{A}_i) - A_i] \\ &= Y_iB_{A_i} \end{aligned} \tag{25}$$

where B_{A_i} is the acreage bias for this *ith* state. The quantities Y_i and B_{A_i} are unknown, but an estimate, $\hat{B}_{P_i}$ can be obtained by using the estimates for Y_i and B_{A_i}. Thus,

$$\hat{B}_{P_i} = \hat{Y}_i \hat{B}_{A_i} \tag{26}$$

The variance of $\hat{B}_{P_i}$ is given by

$$\mathrm{Var}\left(\hat{B}_{P_i}\right) = Y_i^2 \ \mathrm{Var}\left(\hat{B}_{A_i}\right) + E^2\left(\hat{B}_{A_i}\right) \mathrm{Var}(\hat{Y}_i) + \mathrm{Var}\left(\hat{B}_{A_i}\right) \mathrm{Var}(\hat{Y}_i) \tag{27}$$

and estimated by

$$\hat{\mathrm{Var}}\left(\hat{B}_{P_i}\right) = \hat{Y}_i^2 \ \hat{\mathrm{Var}}\left(\hat{B}_{A_i}\right) + \hat{B}_{A_i}^2 \ \hat{\mathrm{Var}}(\hat{Y}_i) - \hat{\mathrm{Var}}\left(\hat{B}_{A_i}\right) \hat{\mathrm{Var}}(\hat{Y}_i) \tag{28}$$

For the nine-state level, the production bias estimate $\hat{B}_P$ is

$$\hat{B}_P = \sum \hat{B}_{P_i} = \sum \hat{Y}_i \hat{B}_{A_i} \tag{29}$$

and the estimate of its variance is $\sum \hat{\mathrm{Var}}\left(\hat{B}_{P_i}\right)$. The relative bias of the production estimate $R(\hat{B}_P)$ is estimated by expressing the production bias as a percentage of the LACIE production estimate; that is, by

$$R(\hat{B}_P) = \frac{\sum \hat{Y}_i \hat{B}_A}{\sum \hat{A}_i \hat{Y}_i} \times 100\% \tag{30}$$

6.2.4.7 <u>Effects of Errors in Acreage, Yield, Sampling, and Classification on the Production Variance</u>

The production variance consists of two major error components: acreage and yield. The acreage error may be further subdivided into sampling and classification errors. The effect of a particular error is determined by the reduction in the production variance estimate when the error is omitted from the calculation

of that estimate. If there is only one yield stratum in a zone (state), the production variance is calculated at the zone level and aggregated to higher levels. If a zone contains more than one yield stratum, it is subdivided into pseudozones, which are the intersections of the zone with the various yield strata. The production variance estimate is then calculated at the pseudozone level and aggregated to the zone and higher levels.

Suppose the zone consists of H pseudozones, $G_1, G_2, \cdots, G_H$, with acreage estimates $A_{Z1}, A_{Z2}, \cdots, A_{ZH}$ and yield predictions $Y_{Z1}, Y_{Z2}, \cdots, Y_{ZH}$, respectively. Then the estimate of the production variance at the zone level is given by the following equation, which also appears in appendix B of the CAS Requirement Document.

$$S_Z^2 = \sum_{i=1}^{H} \left(V_{Zi}^2 Y_{Zi}^2 + U_{Zi}^2 A_{Zi}^2 - V_{Zi}^2 U_{Zi}^2 \right) + 2 \sum_{i=2}^{H} \sum_{\ell=1}^{i-1} Y_{Zi} Y_{Z\ell} \left(\sum_{j \in G_i} \sum_{k \in G_\ell} \Psi_{jk} \right) \tag{31}$$

where

U_{Zi}^2 = the estimate of the yield variance for the *ith* pseudozone

V_{Zi}^2 = the area variance estimate for the *ith* pseudozone

Ψ_{jk} = the estimated covariance between A_j in G_i and A_k in G_ℓ

In order to determine the production variance without a given error term, equation (31) must be rederived with that term omitted. Let S_{ZA}^2, S_{ZY}^2, S_{ZS}^2, and S_{ZC}^2 be the state production variances without acreage, yield, sampling, and classification errors, respectively. One obtains the following expressions for these quantities.

$$S_{ZA}^{2} = \sum_{i=1}^{H} \left(U_{Zi}^{2} A_{Zi}^{2} - V_{Zi}^{2} U_{Zi}^{2}\right) \tag{32}$$

$$S_{ZY}^{2} = \sum_{i=1}^{H} \left(V_{Zi}^{2} Y_{Zi}^{2} - V_{Zi}^{2} U_{Zi}^{2}\right)$$

$$+ 2 \sum_{i=2}^{H} \sum_{\ell=1}^{i-1} Y_{Zi} Y_{Z\ell} \left(\sum_{j \epsilon G_i} \sum_{k \epsilon G_\ell} \Psi_{jk}\right) \tag{33}$$

$$S_{ZS}^{2} = \sum_{i=1}^{H} \left[(1 - \hat{\rho}) V_{Zi}^{2} Y_{Zi}^{2} + U_{Zi}^{2} A_{Zi}^{2} - (1 - \hat{\rho}) V_{Zi}^{2} U_{Zi}^{2}\right]$$

$$+ 2 \sum_{i=2}^{H} \sum_{\ell=1}^{i-1} Y_{Zi} Y_{Z\ell} \left((1 - \hat{\rho}) \sum_{j \epsilon G_i} \sum_{k \epsilon G_\ell} \Psi_{jk}\right) \tag{34}$$

$$S_{ZC}^{2} = \sum_{i=1}^{H} \left(\hat{\rho} V_{Zi}^{2} Y_{Zi}^{2} + U_{Zi}^{2} A_{Zi}^{2} - \hat{\rho} V_{Zi}^{2} U_{Zi}^{2}\right)$$

$$+ 2 \sum_{i=2}^{H} \sum_{\ell=1}^{i-1} Y_{Zi} Y_{Z\ell} \left(\hat{\rho} \sum_{j \epsilon G_i} \sum_{k \epsilon G_\ell} \Psi_{jk}\right) \tag{35}$$

where $\hat{\rho}$ is defined on page C-7 of this report.

Let S_{rA}^2, S_{rY}^2, S_{rS}^2, and S_{rC}^2 be the regional-level production variance estimates without acreage, yield, sampling, and classification errors, respectively. These estimates can be obtained from the following expressions.

$$S_{rA}^{2} = \sum_{Z=1}^{R} S_{ZA}^{2} + \sum_{\substack{Z=1 \\ Z \neq Z'}}^{R} \sum_{Z'=1}^{R} S_{rZZ'} \tag{36}$$

$$S_{rY}^2 = \sum_{Z=1}^{R} S_{ZY}^2 \tag{37}$$

$$S_{rS}^2 = \sum_{Z=1}^{R} S_{ZS}^2 + \sum_{\substack{Z=1 \\ Z \neq Z'}}^{R} \sum_{Z'=1}^{R} S_{rZZ'} \tag{38}$$

$$S_{ZC}^2 = \sum_{Z=1}^{R} S_{ZC}^2 + \sum_{\substack{Z=1 \\ Z \neq Z'}}^{R} \sum_{Z'=1}^{R} S_{rZZ'} \tag{39}$$

Here R is the total number of zones in the region and $S_{rZZ'} = 0$ if Zth and $Z'th$ zones have no yield strata in common. Otherwise,

$$S_{rZZ'} = \sum_{K=1}^{C} A_{rZK} A_{rZ'K} U_{rK}^2 \tag{40}$$

where

A_{rZK} = the area estimate for the pseudozone corresponding to yield stratum K in zone Z of region r

U_{rK}^2 = the squared prediction error for the Kth yield stratum common to zones Z and Z'

C = the number of yield strata common to the Zth and $Z'th$ zones

The estimates of the corresponding variances for a country are obtained by adding the corresponding estimates for all the regions in the country. These computations assume that the regional production estimates are uncorrelated.

6.2.5 SECOND-ORDER ERROR SOURCE INVESTIGATIONS

Second-order error source investigations are designed to study the dependence of the errors in the LACIE system on various

causal factors. It is hoped that such studies will suggest changes in LACIE procedures that will lead to improved estimates.

The second-order error source investigations have been grouped into two categories: segment level and pixel level. Segment level investigations start from data relating to a whole segment, such as a segment wheat proportion. Pixel level investigations start from data relating to pixels, or collections of pixels, such as dots or clusters.

6.2.5.1 Segment-Level Error Source Evaluations

The following paragraphs describe the segment-level data error source evaluation tasks that are planned during Phase III.

6.2.5.1.1 Effect of Bias Correction on Wheat Proportions

The purpose of this task is to determine the error in the LACIE Wheat Proportion Estimate ($\hat{X}$) before and after bias correction has been applied.

To accomplish this, classification and ground-truth data (X) over the blind sites of the USGP are utilized. Comparisons of $\hat{X}$ versus X are made as well as $\hat{X}$ corrected (bias correction applied) versus X. Aggregations are made such that the classification biases and variances can be examined at the state, subregion, and USGP nine-state levels with or without bias correction.

6.2.5.1.2 Effect of Bias Correction on Small Grains Proportion Estimates

This task is similar to that described in section 6.2.5.1.1 except that small grains estimates are evaluated and no aggregations will be performed.

6.2.5.1.3 CAMS Evaluation Code

This task consists of the investigation of the relationship of the CAMS code (describing the quality of the classification) to the proportion error ($\hat{X} - X$) observed over the U.S. blind sites (211). Comparison between the proportion error associated with the categories of CAMS evaluation codes is made and evaluated at state and region levels.

6.2.5.1.4 Acquisition History

The purpose of these evaluations is to see if the observed proportion errors ($\hat{X} - X$) depend on the acquisitions (as represented by the Robertson biostage) during the winter wheat or spring wheat growing seasons. The relationship of error magnitude to acquisition pattern is examined.

The data set for this task includes CAMS classification data and ground-truth data during the 1976-77 growing season over the USGP blind sites recorded by biowindows.

6.2.5.1.5 Haze Effects Evaluations

During Phase III haze effects upon classification are investigated through the acquisition of optical depth measurements with manual radiometers at each ITS at 10-minute intervals from 30 minutes before to 30 minutes after each Landsat 2 overpass of the site. These data shall be collected from February (or end of dormancy) to the end of harvest. The solar radiometer must be calibrated before and after field use and the data must be analyzed with respect to optical depth by November 1, 1977.

In addition to the haze effects study of the above task, during Phase III, three auto tracking solar radiometers are placed in the three supersites (Finney County, Kansas; Hand County, South Dakota; and Williams County, North Dakota) by the field measurements team for use during selected overpasses of Landsat 2.

Optical depth is used and calculated for six channels at rates greater than 1 per second for at least 30 minutes before till 30 minutes after the Landsat 2 overpass. Wind speed and direction, temperature, and relative humidity are also monitored at the same rate over the same period.

Accuracy evaluations of the manually operated solar radiometer involve calculating the mean and standard deviation of the optical depth for each of six channels for the 1-hour observation period, that is correlated with the Landsat overpass, over the 24 USGP ITS's. The correlation between these data and labeling and classification omission and commission errors will be determined.

6.2.5.1.6 Crop Calendar Error Determination

A major reference utilized by analyst-interpreters in their classification procedures is the nominal (mean historical) crop calendar and the adjustable crop calendar (ACC). Since the LACIE ACC provides the latest reference information on the stage of development of wheat in an area being classified and estimated, it is necessary to determine the accuracy of this reference information. This task is designed to determine the accuracy of the ACC estimates of the wheat growth stage throughout the growing season.

The basic data set for these evaluations is the growth-stage data acquired by USDA/ASCS personnel over the 24 U.S. ITS's. These growth-stage data are acquired utilizing the ground-truth periodic observation form presented in appendix E, figure E-2. In addition, growth-stage information is acquired over the following Crop Reporting Districts (CRD's):

a. Texas (1N, 1S, 2N, 2S, 5N, 5S, 8N, and 8S)

b. Kansas (all CRD's)

c. North Dakota (all CRD's)

d. Montana (1, 2, and 3)

The CRD growth-stage data are acquired by USDA/ASCS personnel using the growth-stage reporting forms presented in appendix F.

The USDA/ASCS delivers these data to the DAPTS at JSC 30 days after completion of each applicable growth stage in each CRD stratum or equivalent.

Plots will be made of the ACC outputs (for the ITS's), the mean of the ground observations of wheat growth stages, and the nominal crop calendar. Confidence estimates will be made based on the distribution of the ITS ground-truth observations, and it will be determined if the ACC results fall within these limits. The relationship of the crop calendar information to known episodic events of the current year, such as drought, is investigated and reported throughout Phase III AA, along with the assessment of the accuracy of the ACC.

6.2.5.2 <u>Pixel-Level Error Source Investigations on Blind Sites</u>

These investigations will provide a method of more accurately assessing the labeling, clustering, and classification performance during Phase III. They are based on a new procedure for processing the ground-truth data which is described in appendix H. In this procedure, a "ground-truth tape" is produced in which the ground-truth data are presented as an image, analogous to the Landsat imagery and to the cluster and classification maps produced by CAMS. Each subclass in the ground-truth data has its own assigned grey-scale level on the ground-truth tape. The subclasses used are shown in table 6-2. In addition, certain individual fields (the 15 fields described in section 6.1.5 for blind sites and the 50 fields described in section 6.1.7 for

TABLE 6-2.— SUBCLASSES USED IN AA INVESTIGATIONS AND THE CORRESPONDING GREY-SCALE LEVELS ON THE GROUND-TRUTH TAPE.

Subclass	Grey-scale level
Fields 1 to 80	1 to 80
Alfalfa	90
Beans	91
Corn	92
Safflower	93
Sunflower	94
Sudan grass	95
Sorghum	96
Soybeans	97
Sugar beets	98
Winter wheat	99
Spring wheat	100
Barley	101
Rye	102
Flax	103
Oats	104
Grass	105
Hay	106
Pasture	107
Trees	108
Same as 90-108 except:	
Harvested	115-133
Abandoned	140-158
Strip fallow	165-183
Strip fallow harvested	190-208
Strip fallow abandoned	215-233
Water	240
Homestead	250
Idle cropland stubble	251
Idle cropland cover crop	252
Idle cropland residue	253
Idle cropland fallow	254

ITS's) also have their own assigned grey-scale levels. The image on the ground-truth tape is registered to the corresponding Landsat image. However, the data on the ground-truth tape are at a finer resolution. There are 6 subpixels on the ground-truth image for each pixel on the Landsat image.

In the pixel-level investigations the ground-truth data are compared with the AI dot data and with the cluster and classification maps produced by CAMS. This allows a determination of the actual composition (in terms of ground-truth classes) of each dot, of each cluster on the cluster map, and of each class on the class map. Computer programs which do this automatically are being developed.

The schedule for the processing of the pixel-level data is shown in figure 6-3. It will be seen that the data for Nebraska and Texas will not be processed in time to be included in the final report.

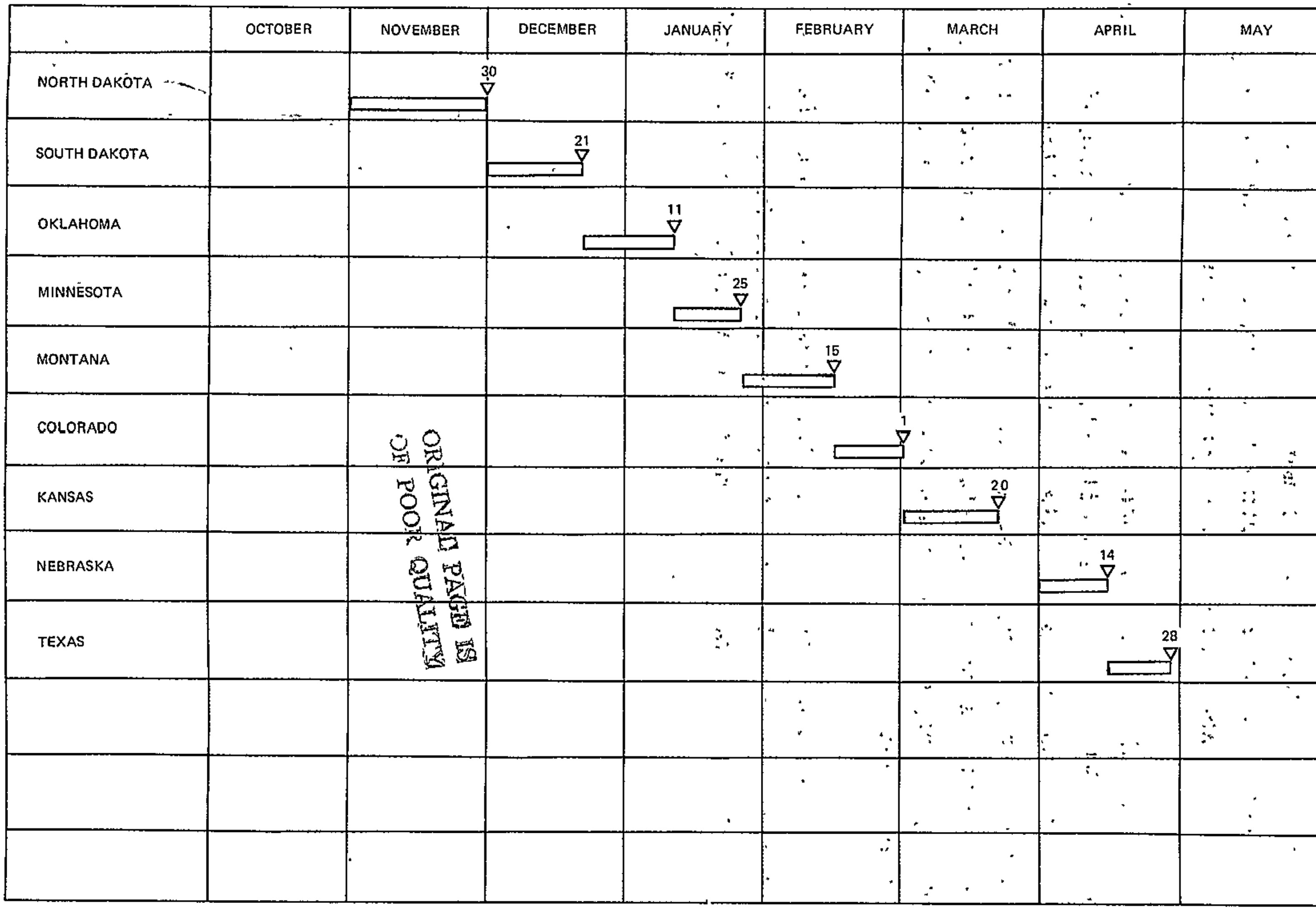


Figure 6-3.— Pixel-level data processing schedule.

6.2.5.2.1 Dot Labeling

The data set is all the analyst labels for the blind sites, starting April 1, 1977, with the small fields bias correction sheets and the Procedure 1 starting and bias correction labels recorded in the CAMS packets since June 1, 1977. These data will have to be key punched by AA for their analyses.

In order to investigate dot labeling, the composition of each dot is obtained first. This consists of determining the representation of the various ground-truth classes (table 6-2) among the 6 subpixels on the ground-truth tape corresponding to each dot. The composition of each dot will be printed out since it is of interest to study the extent to which the dots represent mixtures of different subclasses. Each dot is then labeled with the label of the subclass having the largest representation among the 6 subpixels corresponding to that dot on the ground-truth tape. In what follows, the term "name" will be used to indicate the ground-truth label given to a dot in this manner. This is to distinguish it from the "label" given to it by the analyst-interpreter. It is possible for more than one name to be assigned if two or more subclasses have the same representation. The names are denoted C_n^m where C_n denotes the nth subclass and the superscript m denotes the number of subpixels of this class in the dot. The quantity m/6 is called the "purity" of the dot since it is a measure of the extent to which the dot is composed of a single subclass. The term "set" will be used to indicate the collection of dots named C_n^m.

Each dot is also given a class name (as distinguished from its subclass name). The classes are those used to label the dots by the analyst. For Procedure 1, they are expected to be the following: spring grains (SG), spring wheat (SW), winter grains (WG), winter wheat (WW), grains (G), wheat (W), other (O), and a class denoted X which consists of dots that fell on cloud or

cloud shadow and, therefore, were "unidentifiable." These classes will be denoted U_i, $i = 1, 8$ (i.e., U_1 = SG, U_2 = SW, etc.). The class names are obtained in the same way as the subclass names and are denoted U_i^m, where m is the number of subpixels of class U_i in the dot.

Dot labeling accuracy is studied by estimating two confusion matrices — one for classes and one for subclasses. The class confusion matrix element $P(U_j|U_i^m)$ is the probability that a dot with the name U_i^m will be labeled U_j by the analyst. The elements for which $i = j$ give the probability of correct labeling for each set U_i^m. For $m = 6$, one obtains the probabilities of correct labeling for pure dots; for $m < 6$, one obtains the corresponding probabilities for mixed dots. The other matrix elements give the probabilities for confusion with the various other classes. The class confusion matrix shows how well the analyst-interpreter is performing at labeling pure and mixed pixels of the major classes.

The subclass confusion matrix element $P(U_j|C_n^m)$ is the probability that a dot with name C_n^m will be labeled U_j by an analyst. A particular subclass $C_{n'}$ corresponds to a particular class $U_{j'}$, and the probability of correct labeling for the dots named $C_{n'}^m$ is given by $P(U_{j'} \mid C_{n'}^m)$. The other elements give the probabilities for confusion with the various other classes. The subclass confusion matrix shows how well the analyst-interpreter is performing on the various subclasses and will indicate any classes that pose a particular problem.

Labeling accuracy will be a function of many factors; a partial list of these is given in table 6-3. The effect of these factors on labeling accuracy will be evaluated by AA to the extent permitted by available resources.

TABLE 6-3.– FACTORS AFFECTING LABELING AND CLASSIFICATION ACCURACY

Factor	Description
1	Crop calendar error
2	Acquisition history (including biostage)
	Used by analyst
	Used in classification
3	Percentage of wheat in segment
4	Percentage of small grains in segment
5	Percentage of other crops and idle cropland in segment
6	Percentage of pasture in segment
7	Percentage of grasses in segment
8	Percentage of irrigation in segment
9	Soil type
10	Elevation
11	Crop moisture index
12	Palmer index
13	CAMS evaluation code
14	Analyst
15	Items 3 to 8, historical—
16	Field size
17	Amount of grazing
18	CAMS PCC[a] for type 1 dots
19	CAMS PCC for type 2 dots

[a]The CAMS PCC refers to the agreement between the analyst dot labels and the assignment of class by the classifier.

Finally, a study will be made to determine whether the probability of a dot being correctly labeled is higher if the analyst label agrees with the classifier label for that dot.

6.2.5.2.2 Clustering

Three aspects of clustering will be studied: Cluster composition and purity, cluster labeling accuracy, and the cluster confusion matrix. The data set to be used is all blind site cluster maps from April 1, 1977, to the end of Phase III. These data currently are in the form of DTRM tapes which AA is receiving from ISSRS.

Cluster composition is the percent of the subpixels in a given cluster that belong to each of the major classes. This is determined by comparing the cluster map with the image on the ground-truth tape. The major classes are SG, WG, G, O, and Y, where Y consists of both designated other (DO) and designated unidentifiable (DU) areas. The set of all these classes will be denoted V_j. The "purity" of a cluster is the percentage of the total number of subpixels in the cluster that belong to the class with the largest representation. The composition and purity of clusters are of interest since they indicate how well the clustering algorithm is able to separate the classes into relatively "pure" clusters. These quantities will be studied as a function of segment, region, and acquisition history.

Cluster labeling will be studied by first naming each cluster with the name of the class V_j having the largest representation of subpixels in the cluster. The cluster is correctly labeled if the label given by the labeling logic corresponds to this name. In the case of nearest-neighbor labeling logic, an incorrect label may result from analyst-interpreter mislabeling of the dot used to label the cluster or poor performance by the labeling logic. If the identity of the dots, which were

used to label each cluster, can be determined, these two sources of error will be studied separately. Cluster labeling accuracy will be studied as a function of cluster purity, acquisition history, segment, and region.

Two confusion matrices are estimated for clusters — a class confusion matrix and a subclass confusion matrix. The clustering confusion matrices will be studied as a function of segment, region, and acquisition history.

The class confusion matrix element $P(V_j|V_i)$ is the probability that a subpixel of class V_i will appear in a cluster labeled V_j. For $i = j$, this gives the probability of correct clustering; for $i \neq j$, it gives the probabilities of the clustering algorithm assigning a subpixel of name V_j to the various confusion classes.

The subclass confusion matrix is defined in an analagous fashion. The matrix element $P(V_j|C_n)$ is the probability that a subpixel of subclass C_n is assigned to a cluster labeled V_j. If C_n is a subclass of V_j, the assignment is correct; otherwise, it is not. The study of the $P(V_j|C_n)$ allows one to determine if any particular subclass or group of subclasses is not getting properly clustered.

6.2.5.2.3 Classification Accuracy

Classification will be studied by estimating the classification confusion matrices for classes and subclasses. The data set to be used is the classification files as the DTRM tapes described in the previous section. For classification, the classes are SG, WG, G, O, X, and T, where T indicates pixels which have been thresholded by the classifier. They are denoted W_j, $j = 1, 6$. The subclasses are the same as for dot labeling and clustering.

The class confusion matrix element $P(W_j|W_i)$ is the probability that a subpixel of class W_i is classified as class W_j. For $i = j$, this gives the probability of correct classification; for $i \neq j$, it gives the probabilities of misclassification into the various other classes.

The subclass confusion matrix element $P(W_j|C_n)$ is the probability that a subpixel of class C_n is classified as W_j. If C_n is a subclass of W_j, the classification is correct; otherwise, it is not. The study of $P(W_j|C_n)$ allows one to determine if any particular subclass or group of subclasses is not getting properly classified.

To the extent permitted by available resources, the effect of the various factors in table 6-3 on the classification confusion matrix elements will be evaluated.

6.2.5.2.4 Effect of Crop Height and Ground Cover on Classification Accuracy

This study will use the crop height and ground-cover data acquired every 18 days for 15 selected wheat fields in each blind site. The computer program SPECTL (appendix H) will compute the probability of correct classification for each of these fields and this will be plotted as a function of crop height, Robertson biostage, ground cover, and "green number." Means and other relevant statistics will be calculated at the segment, state, and regional levels.

6.2.5.3 Pixel-Level Investigations for ITS's

CAMS will process the 24 ITS's in the United States in the same manner as regular LACIE segments are processed. The objective in this analysis is to determine labeling and classification accuracies. For winter wheat, this will be done for all three procedures used to determine winter wheat proportions, namely

the Phase II procedure, the small fields procedure, and Procedure 1. For spring wheat, only Procedure 1 will be investigated because it is the only procedure to be used with spring wheat.

The ground-truth data consist of two complete inventories of the test site, one in the fall of 1976 and the other in the spring of 1977. Between these periods, updates are provided every 18 days, coincident with the satellite overpass, for a subset of the fields in the test site. This subset is called the 18-day fields. The performance evaluation analysts will use these updates to correct the ground-truth designations for the 18-day fields on the ground-truth map and keep them current. Also, on the basis of the information available for the 18-day fields, the analysts will interpret on a current PFC image to update the designations for all other fields in the test site; this will provide a complete, wall-to-wall ground-truth image for the test site. A similar photointerpretation will be performed to obtain ground-truth designations for the fields outside the test site but within the LACIE segment containing the test site.

In the Phase II and the small fields procedures, labeling and classification accuracy will be determined for each classification for the test site, and a set of 30 fields chosen at random from all the fields in the test site.

For Procedure 1, labeling accuracy and proportion error will be determined for each classification, both for the area within the test site and for the whole LACIE segment.

The "true" proportion of the area within the test site will be determined by adding the acreages of the fields in the site. For the whole segment, the "true" proportion will be determined by applying a bias correction, based on ground-truth labels, to the machine proportions of the segment.

The results of these investigations will be reported in the interim reports and at least one ITS story will be presented in the quick-look following the CAS CMR. Thus, the ITS can fulfill a role not possible by the blind sites — illustration of accuracy using annotated imagery.

This effort starts February 1977 when the first inventories are received and ends in October when the harvest inventory is received.

6.2.5.4 Second-Order Investigations on the IMAGE 100 System

These investigations are designed to evaluate results obtained with the IMAGE 100 Hybrid Procedure 1 operational system.

6.2.5.4.1 Study of Proportion Error

The data set will consist of 24 ITS's in the United States and 10 ITS's in Canada. The IMAGE 100 proportions will be determined for the test sites (which are smaller than the whole segment) and compared with the corresponding ground-truth proportions determined in the manner described in section 6.2.5.3. The method used to make the comparison is described in section 6.2.4.1.

6.2.5.4.2 Effects of AI, Acquisition History, and Bias Correction on Proportion Estimation Error

The IMAGE 100 processor and data from eight U.S. blind sites will be used in an experiment wherein each site will be analyzed by three AI's to give a Procedure 1 "raw" and a "bias-corrected" estimate of the proportion of small grains in each segment. The segments will be of two types; namely, those having acquisitions in all four biophases and those having only early season acquisitions. The segments will be chosen at random from the blind sites for which detailed ground-truth data are available.

The objectives of the experiment are: (1) to evaluate the performance of Procedure 1 in terms of absolute proportion estimation error and its repeatibility over AI's, (2) to make comparisons between "bias-corrected" and "raw" Procedure 1 estimates, and (3) to determine if the performance is better when acquisitions from all biostages are used than when only the early season acquisition is used.

Analysis of variance will be used to test for the effects of AI's, time (i.e., early season versus all acquisitions), method (raw versus bias correction estimates), and their interactions.

6.2.6 ACCURACY ASSESSMENT EVALUATION OF THE NEW SAMPLE STRATEGY

During LACIE Phase III a new sample strategy is being tested over Kansas and North Dakota. AA evaluations are designed to examine how well the new sampling strategy is doing over these areas through comparisons between the old and new strategies. The evaluations utilize (a) blind site LACIE classification and ground-truth data collected during the 1976-1977 growing season over Kansas and North Dakota, (b) 1976-1977 LACIE and USDA/SRS production, area, and yield estimates over these same states, and (c) 1974 USDA/SRS historical county data also for Kansas and North Dakota.

The AA evaluations of the new sample strategy consist of:

a. Subtask 1 making comparisons (relative differences and CV's) between the 1977 LACIE estimates of production, area, and yield over Kansas and North Dakota from the old and new sample strategies, and estimates made by USDA/SRS. Figure 6-4 further defines these comparisons.

b. Subtask 2 determining and comparing aggregated acreage bias for old and new sample strategies over Kansas (winter wheat)

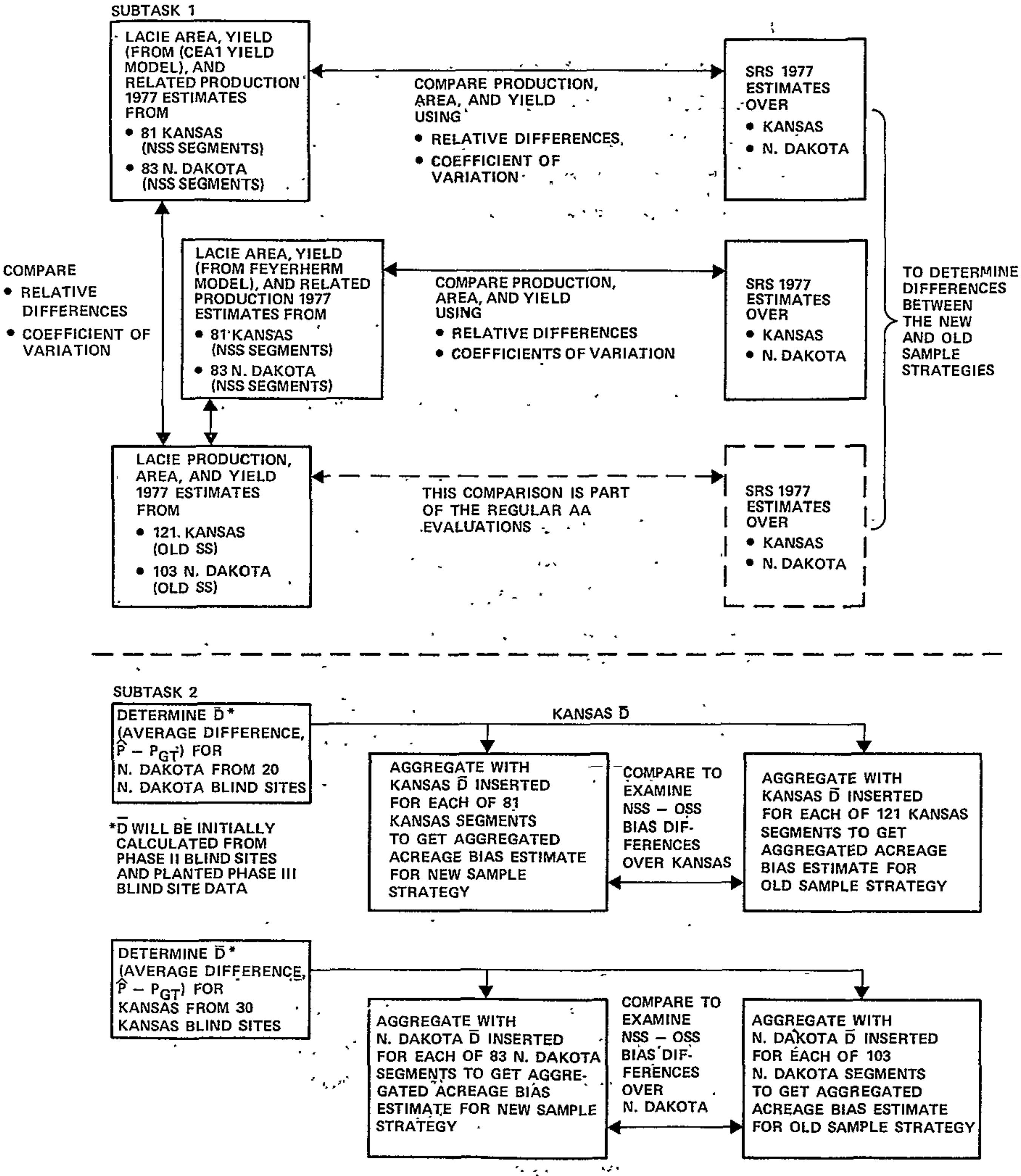


Figure 6-4.— AA new sample strategy (NSS) evaluations.

and North Dakota (spring wheat). Figure 6-4 also provides an additional description of these evaluations.

c. Subtask 3 performing regression analysis using 1977 blind site classification data, ground-truth data, and USDA/SRS 1974 historical census data to obtain three basic regressions.

   - True segment proportion vs. historical stratum proportion
   - LACIE segment proportion vs. ground truth segment proportion
   - LACIE segment proportion vs. historical stratum proportion

Using the results of these regressions, a sensitivity analysis is performed in which the acreage error is quantified (a) with both the sampling and the classification errors, (b) with the sampling error and no classification error, and (c) with the classification error and no sampling error. Then the reduction in error from (a)-(b) and (a)-(c) is examined to establish which, if either, is contributing more error to the acreage estimate.

6.3 AA REPORTING

There are three types of AA reports: Special management briefings, monthly quick-look reports, and interim and final reports.

6.3.1 AA SPECIAL MANAGEMENT BRIEFINGS

As required to support LACIE reporting functions, special briefings are prepared for LACIE project management relative to the current status of LACIE operational data development, particularly with respect to special problems that could affect the accuracy of final LACIE at-harvest estimates of wheat acreage, yield, and production in the USGP.

In addition, AA personnel support LACIE project management requirements to brief NASA, USDA, and NOAA upper-level management on the status of LACIE outputs and the progress being made toward satisfying the 90-90 criterion.

6.3.2 AA MONTHLY QUICK-LOOK REPORTS

The quick-look reports contain an evaluation by AA of the LACIE estimates reported in the CMR's and the CAR. The quick-look reports are released one week following the release of a CMR or a CAR. The CMR's and CAR's contain the official LACIE estimates of wheat production, area, and yield, and the corresponding statistics. The true wheat production, area, and yield for the particular region or country are, of course, unknown. Therefore, to ascertain the accuracy of the LACIE estimates, comparisons are made with a reference standard. In the United States, the reference standard consists of the most recent (at the time of the comparison) estimates released by the USDA/SRS. In foreign countries, the reference consists of the most recent estimates released by the USDA/FAS.

In addition, the quick-look reports will contain (a) significance tests of no difference (between the LACIE estimates and the reference standard) at the region or country level; (b) results of blind site analyses of proportion estimation error; (c) classification bias aggregated to the regional level; and (d) within stratum acreage variance due to classification and sampling errors.

6.3.3 AA INTERIM AND FINAL REPORTS

The interim reports are released at regular intervals throughout the crop season. They contain the results of the previous quick-look reports, and the results of all other AA investigations up to the time each report is written.

Each interim report is built up from the previous one by including data that became available during the interim period. Technical comments on each report are solicited from a variety of sources and are used to upgrade subsequent reports. Early and mid-season evaluations are made in the first and second interim

reports; late season and at-harvest evaluations are made in the third and fourth interim reports.

The fourth interim report also serves as a draft for the final report, which contains material similar to the interim reports but covers the entire year.

The planned schedule for release of the LACIE interim AA reports and the final Phase III AA report is as follows:

a. First interim report - May 15, 1977

b. Second interim report - August 15, 1977

c. Third interim report - November 15, 1977

d. Final interim report - March 1, 1978

e. Final Phase III AA report - June 1, 1978

The planned contents for the four interim reports and the final report are shown in figure 6-5. The AA interim reports require approval by the AA manager and the Chief of the RTEB of NASA/JSC prior to their release for LACIE project review and evaluation.

The final LACIE Phase III AA report requires review and approval of the following persons prior to its release for distribution:

a. D. E. Pitts, Manager, LACIE AA, NASA/JSC

b. J. D. Erickson, Chief, RTEB, NASA/JSC

c. J. L. Dragg, Chief, Applications System Verification Branch, NASA/JSC

d. F. G. Hall, LACIE Project Scientist, NASA/JSC

e. J. Hill, Assistant Deputy Project Manager for LACIE, NOAA

f. J. Murphy, Assistant Deputy Project Manager for LACIE, USDA

FIRST INTERIM REPORT (MAY 15, 1977)	SECOND INTERIM REPORT (AUGUST 15, 1977)	THIRD INTERIM REPORT (NOVEMBER 15, 1977)	FOURTH INTERIM REPORT (MARCH 1, 1978)	
DATA SETS • U S CMR'S (LACIE ESTIMATES – FEBRUARY, MARCH, AND APRIL) • GROUND TRUTH (PLANTED) – 45 FALL WW BLIND SETS – USSGP • 1977 SRS STATE ESTIMATES	DATA SETS • U.S CMR'S (LACIE ESTIMATES – FEBRUARY, APRIL, MAY, JUNE, AND JULY) • GROUND TRUTH (PLANTED) – 45 FALL WW BLIND SITES • 1977 SRS STATE ESTIMATES	DATA SETS • U.S AND U S S R CMR'S (LACIE ESTIMATES – FEBRUARY AND OCTOBER) • GROUND TRUTH (HARVESTED) – 84 WW BLIND SITES – 59 SW BLIND SITES – 18 ITS • 1977 SRS STATE ESTIMATES • 1977 U.S S.R ESTIMATES	DATA SETS • U S. AND U S S R CMR'S (THROUGH END OF YEAR) • CAR • GROUND TRUTH (HARVESTED) – 135 WW BLIND SITES – 65 SW BLIND SITES • 1977 SRS STATE ESTIMATES • 1977 U S S R ESTIMATES	REPORT CONTENTS ARE THE SAME AS FOR THE FOURTH INTERIM REPORT
EVALUATIONS • LACIE VS SRS ESTIMATES COMPARISONS – 1977 ESTIMATES OVER 5 STATE WW REGION – 1977 LACIE COMPARISONS WITH 1976 LACIE	EVALUATIONS • COMPARISONS – 1977 LACIE VS 1977 SRS ESTIMATES OVER 7-STATE WW REGION – 1977 LACIE COMPARISON WITH 1976 LACIE	EVALUATIONS • COMPARISONS – 1977 LACIE VS 1977 SRS ESTIMATES • ALL STATES IN USGP • ALL REGIONS IN USGP • COUNTY LEVEL IN U.S S R – 1977 LACIE COMPARISON WITH 1976 LACIE	EVALUATIONS REPORT CONTENTS ARE SAME AS FOR THE THIRD INTERIM REPORT	
• 90 90 CRITERION EVALUATIONS OVER 5 STATE WW AREA	• 90 90 CRITERION EVALUATIONS OVER 5 STATE WW AREA	• 90 90 CRITERION EVALUATION OVER 9 STATES	REPORT CONTENTS ARE THE SAME AS FOR THE THIRD INTERIM REPORT	
• FIRST ORDER ERROR SOURCE EVALUATION FOR 5 STATE USSGP – WHEAT (P_W) PROPORTION ERROR – CLASSIFICATION VARIANCE – SAMPLING VARIANCE	• FIRST ORDER ERROR SOURCE EVALUATIONS FOR 5 STATE USSGP – WHEAT (P_W) PROPORTION ERROR – CLASSIFICATION VARIANCE – SAMPLING VARIANCE	• FIRST ORDER ERROR SOURCE EVALUATIONS FOR USGP – WHEAT PROPORTION ERROR – SMALL GRAIN PROPORTION ERROR – CLASSIFICATION BIAS – CLASSIFICATION VARIANCE – SAMPLING VARIANCE –PROCEDURE 1 HYBRID ACCURACY	• FIRST ORDER ERROR SOURCE EVALUATIONS FOR USGP REPORT CONTENTS ARE THE SAME AS FOR THE THIRD INTERIM REPORT	
• SECOND ORDER SEGMENT LEVEL EVALUATIONS – SPECIAL STUDIES	• SECOND ORDER SEGMENT LEVEL EVALUATIONS – THRESHOLDING EFFECTS – SPECIAL STUDIES	• SECOND ORDER SEGMENT LEVEL EVALUATIONS – CAMS EVALUATION CODE – CROP CALENDAR ERROR – PROCEDURE 1 HYBRID ACCURACY	• SECOND ORDER SEGMENT LEVEL EVALUATIONS – CAMS EVALUATION CODE – HAZE EFFECTS – PROPORTION ERROR VS BIOSTAGE – CROP CALENDAR ERROR – PROCEDURE 1 HYBRID ACCURACY	
• SECOND ORDER PIXEL LEVEL EVALUATIONS NO PIXEL LEVEL EVALUATIONS IN THIS REPORT	• SECOND ORDER PIXEL LEVEL EVALUATIONS OF LABELING, CLASSIFICATION, AND CLUSTERING ACCURACY NO PIXEL LEVEL EVALUATIONS IN THIS REPORT	• SECOND ORDER PIXEL LEVEL EVALUATIONS OF LABELING, CLASSIFICATION AND CLUSTERING ACCURACY – 20 BLIND SITES – WHEAT VS OTHER – SMALL GRAIN VS OTHER – BOUNDARY PIXEL LABELING ACCURACY – 15 FIELD ANALYSIS – PROCEDURE 1 HYBRID ACCURACY	• SECOND ORDER PIXEL LEVEL EVALUATIONS OF LABELING CLASSIFICATION, AND CLUSTERING ACCURACY REPORT CONTENTS ARE THE SAME AS FOR THE THIRD INTERIM REPORT.	

Figure 6-5.– Summary of Phase III AA interim and final reports.

g. R. B. Erb, Chief, LACIE Project Office, NASA/JSC

h. R. B. MacDonald, Chief, Earth Observations Division, NASA/JSC

PRECEDING PAGE BLANK NOT FILMED

APPENDIX A

ACCURACY ASSESSMENT TEAM

APPENDIX A

ACCURACY ASSESSMENT TEAM

1. D. E. Pitts, Chairman, NASA/SF3
2. R. M. Bizzell, NASA/SF4
3. A. G. Houston, NASA/SF3
4. J. H. Carney, NASA/SF4
5. A. H. Feiveson, NASA/SF3
6. A. D. Frank, USDA
7. L. O. Lovfald, USDA
8. R. L. Patterson, NASA/SF4
9. R. G. Stuff, NASA/SF3
10. B. L. Carroll, LEC
11. R. S. Chhikara, LEC
12. E. M. Hsu, LEC
13. C. J. Liszcz, LEC
14. M. L. Mannen, LEC
15. J. F. Potter, LEC
16. F. W. Solomon, LEC
17. Duane Marquis, USDA

APPENDIX B

LACIE PHASE III BLIND SITES, TEST SITES, AND INTENSIVE TEST SITES

APPENDIX B

LACIE PHASE III BLIND SITES, TEST SITES, AND INTENSIVE TEST SITES

The LACIE ITS and blind-site selections are given in the following tables.

Table	Sites
B-1	LACIE Phase III U.S. ITS's
B-2	LACIE Phase III Canadian ITS's
B-3	Phase III Test Site and Blind-Site Distribution

TABLE B-1.— LACIE PHASE III U.S. INTENSIVE TEST SITES

Segment	State	County	Center coordinates		Site size		Wheat type*	Acquired as
			Lat., N	Long., W	N. mi.	Km		
1961	Kansas	Morton	37°16'00"	101°54'00"	5×6	9×11	WW	WW
1962	Kansas	Saline	38°41'48"	97°28'24"	3×3	5.5×5.5	WW	WW
1963	Kansas	Rice	38°17'00"	98°12'42"	3×3	5.5×5.5	WW	WW
1964	Kansas	Ellis	38°50'06"	99°13'00"	3×3	5.5×5.5	WW	WW
1988	Kansas	Finney	38°10'12"	100°43'12"	5×6	9×11	WW	WW
1965	N. Dak.	Burke	48°53'12"	102°10'00"	5×6	9×11	SW	SW
1966	N. Dak.	Williams	48°19'12"	103°24'42"	5×6	9×11	SW	SW
1967	N. Dak.	Divide	48°53'36"	103°10'54"	2×10	3.7×18.5	SW	SW
1969	Montana	Toole	48°53'00"	111°46'36"	2×10	3.7×18.5	S&WW	SW
1970	Montana	Liberty	48°44'00"	110°51'00"	2×10	3.7×18.5	S&WW	SW
1971	Montana	Hill	48°42'00"	109°55'00"	2×6	3.7×11	S&WW	SW
1973	Wash.	Whitman 2	46°50'24"	117°48'18"	3×3	5.5×5.5	S&WW	WW
1975	Idaho	Oneida	42°04'30"	112°29'30"	3×3	5.5×5.5	S&WW	WW
1976	Idaho	Franklin	42°08'00"	111°58'00"	3×3	5.5×5.5	S&WW	WW
1977	Idaho	Bannock	42°56'30"	112°25'50"	3×3	5.5×5.5	S&WW	WW
1978	Texas	Randall	35°09'30"	102°04'24"	3×3	5.5×5.5	WW	WW
1979	Texas	Deaf Smith	34°52'12"	102°22'18"	3×3	5.5×5.5	WW	WW
1980	Texas	Oldham	35°15'00"	102°32'00"	3×3	5.5×5.5	WW	WW
1981	Indiana	shelby	39°27'36"	85°47'12"	3×3	5.5×5.5	WW	WW
1982	Indiana	Madison	40°13'30"	85°37'50"	3×3	5.5×5.5	WW	WW
1983	Indiana	Boone	40°05'42"	86°33'90"	3×3	5.5×5.5	WW	WW
1687	S. Dak.	Hand 1	44°35'00"	98°58'00"	5×6	9×11	S&WW	SW
1986	S. Dak.	Hand 2	44°21'00"	98°45'06"	5×6	9×11	S&WW	SW
1987	Minn.	Polk	47°49'00"	96°41'00"	5×6	9×11	SW	SW

*As indicated by ground-truth data:

WW = winter wheat
SW = spring wheat
S&WW = spring and winter wheat

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE B-2.— LACIE PHASE III CANADIAN INTENSIVE TEST SITES

Segment	Province	County	Center coordinates/site size				Wheat type*
			Lat., N	Long., W	S. mi.	Km	
1958	Saskatchewan	Melfort	52°48'	104°44'	2×10	3.2×16	SW
1959	British Columbia	Dawson Creek	55°48'	120°12'	2×10	3.2×16	SW
1984	Saskatchewan	Delisle	51°55'	107°28'	2×10	3.2×16	SW
1985	Saskatchewan	Swift Current	50°19'	107°53'	2×10	3.2×16	SW
1989	Alberta	Lethbridge (Raymond)	49°30'	112°48'	2×10	3.2×16	SW
1990	Manitoba	Stony Mountain	50°04'	97°21'	2×10	3.2×16	SW
1991	Manitoba	Starbuck	49°47'	97°29'	2×10	3.2×16	SW
1992	Alberta	Olds	51°54'	113°32'	2×10	3.2×16	SW
1994	Alberta	Ft. Saskatchewan	53°38'	113°07'	2×10	3.2×16	SW
1995	Manitoba	Altona	49°12'	97°38'	1×5	1.6×8	SW

*As indicated by ground-truth data: SW = spring wheat.

TABLE B-3.– PHASE III TEST SITE AND BLIND-SITE DISTRIBUTION

State	Number of sites
Colorado	15
Kansas	31
Nebraska	26
Oklahoma	20
Texas	18
Minnesota	20
Montana	29
N. Dakota	30
S. Dakota	23
Total	212
Canadian Test Sites	30

PRECEDING PAGE BLANK NOT FILMED

APPENDIX C

CONTRIBUTION OF SAMPLING AND CLASSIFICATION TO ACREAGE ESTIMATION ERROR

APPENDIX C

CONTRIBUTION OF SAMPLING AND CLASSIFICATION TO ACREAGE ESTIMATION ERROR

This appendix describes the calculation of the contribution of sampling and classification errors to the variance of the LACIE production estimate.

C.1 APPROACH

The variance of the LACIE acreage estimate for a large area (e.g., zone) can be written

$$V^2 = \sum_i V_i \sigma_i^2 \tag{1}$$

where σ_i^2 is the variance of the acreage estimate for the *ith* county and V_i is a weight which depends on the size of the county, the number of segments in the county, etc. (Refer to the CAS Requirements Document, appendix B for details.)

The variance σ_i^2 represents a mean-squared deviation between the LACIE estimate for the county wheat proportion and the true county wheat proportion. This variance is caused mainly by two factors: sampling errors and classification errors.

In accuracy assessment, it is desirable to quantify the contribution of each of these error sources to the large area production estimate. The LACIE production estimate depends on acreage and yield estimation errors in a complicated way; hence, it is unrealistic to assume the error in the production estimate can be written as a sum of uncorrelated random variables representing acreage and yield errors. Instead, the effect of a particular error source is measured by the reduction in the LACIE production variance which would be achieved if that source were eliminated.

It will be assumed that the *ith* county acreage error variance σ_i^2 can be written $\sigma_i^2 = \sigma_c^2 + \lambda^2\sigma_s^2$, where σ_c^2 is a contribution due to classification, and $\lambda^2\sigma_s^2$ is a contribution due to sampling. To determine the effect of no classification error, the variance of the LACIE production estimate will be made using $\rho\sigma_i^2$ instead of σ_i^2 where ρ is an estimate of the ratio $\dfrac{\lambda^2\sigma_s^2}{\sigma_c^2 + \lambda^2\sigma_s^2}$. Similarly, the effect of no sampling error is estimated by replacing σ_i^2 by $(1 - \rho)\sigma_i^2$. The following two sections describe the methods employed for estimating sampling and classification variances and the function ρ.

C.2 ACREAGE REGRESSION MODELS

For counties with one sample segment, the LACIE estimate of the *ith* county wheat proportion can be written

$$
\begin{aligned}
y_i &= C_i + (x_i - C_i) + (y_i - x_i) \\
&= C_i + \varepsilon_i + \delta_i \qquad (2)
\end{aligned}
$$

where

y_i = LACIE estimate of the wheat proportion in the sampled segment

C_i = true (current year) proportion of wheat in the county

x_i = true proportion of wheat in the sampled segment

ε_i = sampling error = $x_i - C_i$

δ_i = classification error = $y_i - x_i$

It will be assumed that for some reasonably large area (e.g., a zone) the errors ε_i and δ_i have the following properties:

$$\varepsilon_i \text{ and } \delta_i \text{ are uncorrelated} \tag{3}$$

$$E(\varepsilon_i) = 0 \tag{4}$$

$$E(\delta_i | x_i) = \lambda^* x_i + \theta \tag{5}$$

$$V(\varepsilon_i) = \sigma_s^2 \tag{6}$$

$$V(\delta_i | x_i) = \sigma_c^2 \tag{7}$$

It is also assumed that there is a linear model relating the current year counties proportions, c_i, to the historical proportions which will be denoted by Z_i; i.e.,

$$C_i = \alpha + \beta Z_i + \zeta_i \tag{8}$$

where $E(\zeta_i) = 0$, $V(\zeta_i) = \sigma_H^2$, $\mathrm{Cov}(\zeta_i, \varepsilon_i) = \mathrm{Cov}(\zeta_i, \delta_i) = 0$ and α and β are regression coefficients.

From the above assumptions and definitions, three basic regression models are obtained:

a. True segment proportion vs. historical county proportion — from the definition of ε_i,

$$\begin{aligned} x_i &= C_i + \varepsilon_i \\ &= \alpha + \beta Z_i + \zeta_i + \varepsilon_i \end{aligned} \tag{9}$$

It follows that

$$\left.\begin{aligned} E(x_i) &= \alpha + \beta Z_i \\ V(x_i) &= \sigma_H^2 + \sigma_s^2 \end{aligned}\right\} \tag{10}$$

b. LACIE segment proportion vs. ground truth segment proportion - from the definition of δ_i

$$y_i = x_i + \delta_i.$$

It follows that

$$E(y_i|x_i) = x_i + \lambda^* x_i + \theta$$

$$V(y_i|x_i) = \sigma_c^2$$

Writing $\lambda = 1 + \lambda^*$, one obtains

$$E(y_i|x_i) = \lambda x_i + \theta \tag{11}$$

$$V(y_i|x_i) = \sigma_c^2 \tag{12}$$

c. LACIE segment proportion vs. historical county proportion — from equations (9) through (12),

$$E(y_i) = E_{x_i}\left[E(y_i|x_i)\right] = E_{x_i}(\lambda x_i + \theta) = \lambda(\alpha + \beta Z_i) + \theta \tag{13}$$

$$V(y_i) = E_{x_i}\left[V(y_i|x_i)\right] + V_{x_i}\left[E(y_i|x_i)\right] = \sigma_c^2 + \lambda^2\left(\sigma_H^2 + \sigma_s^2\right) \tag{14}$$

As stated previously, one would like to estimate $\rho = \dfrac{\lambda^2\sigma_s^2}{\sigma_c^2 + \lambda^2\sigma_s^2}$.

None of the three regression models permits an estimate of σ_s^2 separately from σ_H^2; i.e., one can only estimate $\sigma_s^2 + \sigma_H^2$, not σ_s^2 alone. If current year county proportions C_i were available, σ_H^2 could be estimated, but since this is not the case, $\rho^* = \dfrac{\lambda^2\left(\sigma_s^2 + \sigma_H^2\right)}{\sigma_c^2 + \lambda^2\left(\sigma_s^2 + \sigma_H^2\right)}$ will be estimated instead of ρ. If $\sigma_H^2 << \sigma_c^2$ (a reasonable assumption) then $\rho^* \approx \rho$.

C.3 NORMALITY ASSUMPTIONS — MAXIMUM LIKELIHOOD ESTIMATION OF ρ^*

Suppose a given zone has m blind site segments and n ordinary (i.e., not blind site) segments, and let the blind site segments be numbered 1 to m. It is assumed that ground truth wheat proportions $\{x_i\}_{i=1}^{m}$ are available for the blind sites and LACIE estimates $\{y_i\}_{i=1}^{m+n}$ are available for all the segments. It is also assumed that historical wheat proportions $\{Z_i\}_{i=1}^{m+n}$ are available for the counties containing the segments. If $\sigma_H^2 << \sigma_s^2$, so that $\rho \approx \rho^*$ the regression models equations (9 through 14) can be used to obtain

$$E(x_i) = \alpha + \beta Z_i; \; V(x_i) \doteq \sigma_s^2 \qquad i = 1, 2, \cdots, m$$

$$E(y_i|x_i) = \lambda x_i + \theta; \; V(y_i|x_i) = \sigma_c^2 \qquad i = 1, 2, \cdots, m$$

$$E(y_i) = \theta + \lambda\alpha + \lambda\beta Z_i; \; V(y_i) = \lambda^2\sigma_s^2 + \sigma_c^2 \qquad i = m+1, \cdots, m+n$$

If there is one segment per county, then the errors ε_i and δ_i are independent for different values of i, and hence the likelihood function of the sample can be written

$$L = \prod_{i=1}^{m} f(x_i, y_i) \prod_{i=m+1}^{m+n} h(y_i) \tag{15}$$

where $f(x_i, y_i)$ is the joint density of x_i and y_i for $i = 1, m$ and $h(y_i)$ is the density of y_i for $i = m+1, m+n$.

The function $\prod_{i=1}^{m} f(x_i, y_i)$ can be written $\prod_{i=1}^{m} f(x_i, y_i) =$ $\prod_{i=1}^{m} f(y_i|x_i)\, g(x_i)$ where $f(y_i|x_i)$ is the conditional density of y_i given x_i and $g(x_i)$ is the density of x_i.

If normality is assumed, $\prod_{i=1}^{m} f(x_i, y_i) = \prod_{i=1}^{m} \frac{1}{\sigma_c\sqrt{2\pi}}$

$$\exp\left\{-\frac{1}{2\sigma_c^2}\sum_{i=1}^{m}(y_i - \lambda x_i - \theta)^2\right\}\frac{1}{\sigma_s\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma_s^2}\sum_{i=1}^{m}(x_i - \alpha - \beta Z_i)^2\right\}$$

and

$$\prod_{i=m+1}^{m+n} h(y_i) = \frac{1}{\left(\lambda^2\sigma_s^2 + \sigma_c^2\right)^{1/2}\sqrt{2\pi}}\exp\left\{-\frac{1}{2\left(\lambda^2\sigma_s^2 + \sigma_c^2\right)}\sum_{i=m+1}^{m+n}(y_i - \lambda\alpha - \theta - \lambda\beta Z_i)^2\right\}$$

Letting $Q = -2\log L - \log 2\pi$,

$$Q = m\log\sigma_s^2 + m\log\sigma_s^2 + n\log(\sigma_c^2 + \lambda^2\sigma_s^2) + \frac{D_m}{\sigma_c^2} + \frac{T_m}{\sigma_s^2} + \frac{T_n}{\sigma_c^2 + \lambda^2\sigma_s^2} \tag{16}$$

where

$$D_m = \sum_1^m (y_i - \lambda x_i - \theta)^2$$

$$T_m = \sum_1^m (x_i - \alpha - \beta Z_i)^2$$

$$T_n = \sum_{i=m+1}^{m+n}(y_i - \lambda\alpha - \theta - \lambda\beta Z_i)^2$$

One attempts to maximize L by finding a stationary point of Q:

$$-\frac{1}{2}\frac{\partial Q}{\partial \alpha} = \frac{\sum_1^m (x_i - \alpha - \beta Z_i)}{\sigma_s^2} + \frac{\sum_{m+1}^{m+n}\lambda(y_i - \lambda\alpha - \theta - \lambda\beta Z_i)}{\sigma_c^2 + \lambda^2\sigma_s^2} = 0 \tag{17}$$

$$-\frac{1}{2}\frac{\partial Q}{\partial \beta} = \frac{\sum_{1}^{m} z_i(x_i - \alpha - \beta z_i)}{\sigma_s^2} + \frac{\sum_{m+1}^{m+n} \lambda z_i(y_i - \lambda\alpha - \theta - \lambda\beta z_i)}{\sigma_c^2 + \lambda^2\sigma_s^2} = 0 \tag{18}$$

$$-\frac{1}{2}\frac{\partial Q}{\partial \theta} = \frac{\sum_{1}^{m} (y_i - \lambda x_i - \theta)}{\sigma_c^2} + \frac{\sum_{m+1}^{m+n} (y_i - \lambda\alpha - \theta - \lambda\beta z_i)}{\sigma_c^2 + \lambda^2\sigma_s^2} = 0 \tag{19}$$

$$-\frac{1}{2}\frac{\partial Q}{\partial \lambda} = \frac{\sum_{1}^{m} x_i(y_i - \lambda x_i - \theta)}{\sigma_c^2} + \frac{-n\lambda\sigma_s^2 + \sum_{i=m+1}^{m+n} (\beta z_i + \alpha)(y_i - \lambda\alpha - \theta - \lambda\beta z_i)}{\sigma_c^2 + \lambda^2\sigma_s^2}$$

$$+ \frac{\lambda^2\sigma_s^2 T_n}{\left(\sigma_c^2 + \lambda^2\sigma_s^2\right)^2} = 0 \tag{20}$$

$$\frac{\partial Q}{\partial \sigma_c^2} = \frac{m}{\sigma_c^2} + \frac{n}{\lambda^2\sigma_s^2 + \sigma_c^2} - \frac{D_m}{\sigma_c^4} - \frac{T_n}{\left(\lambda^2\sigma_s^2 + \sigma_c^2\right)^2} = 0 \tag{21}$$

$$\frac{\partial Q}{\partial \sigma_s^2} = \frac{m}{\sigma_s^2} + \frac{n\lambda^2}{\lambda^2\sigma_s^2 + \sigma_c^2} - \frac{T_m}{\sigma_s^4} - \frac{T_n\lambda^2}{\left(\sigma_c^2 + \lambda^2\sigma_s^2\right)^2} = 0 \tag{22}$$

Equations (17) through (22) must be solved for the parameters α, β, θ, λ, σ_c^2, and σ_s^2. If $\hat{\alpha}$, $\hat{\beta}$, $\hat{\theta}$, $\hat{\lambda}$, $\hat{\sigma}_c^2$, and $\hat{\sigma}_s^2$ represent the solution to equations (17) through (22), then the invariance theorem for maximum likelihood estimation can be used to obtain

$$\hat{\rho} = \frac{\hat{\lambda}^2\hat{\sigma}_s^2}{\hat{\sigma}_c^2 + \hat{\lambda}^2\hat{\sigma}_s^2} \tag{23}$$

as the maximum likelihood estimate of ρ.

The equations (17) through (22) are nonlinear but can be solved using numerical techniques. Newton's Method was used to solve the equations for this report; i.e., if $u^{(k)}$ is an estimate of the solution vector $u = (\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{\lambda}, \hat{\sigma}_c^2, \hat{\sigma}_s^2)$ at the k*th* step, then

$$u^{(k+1)} = u^{(k)} - F^{-1}f\left(u^{(k)}\right) \tag{24}$$

where $f\left(u^{(k)}\right) = (f_1, \cdots, f_6)^T$ is the vector of the left sides of equations (17) through (22) evaluated at $u^{(k)}$ and $F = (F_{ij}) = \frac{\partial f_i}{\partial u_i}$.

In practice, it was slightly simpler to use the parameter transformations

$$r = \frac{\sigma_s^2}{\lambda^2 \sigma_s^2 + \sigma_c^2} \tag{25}$$

and

$$s = \lambda^2 \sigma_s^2 + \sigma_c^2 \tag{26}$$

and solve for α, β, θ, λ, r, and s. Again, the invariance theorem can be used to give

$$\hat{\rho} = \hat{\lambda}^2 \hat{r} \tag{27}$$

C.4 ACCURACY OF $\hat{\rho}$

Since $\hat{\rho}$ is an extremely complicated function of the data, it is impossible to write down the variance of $\hat{\rho}$ for finite sample sizes m and n. However, the asymptotic variance of $\hat{\rho}$ can be estimated using the information matrix; i.e., if

$$\tau = E\left\{\frac{-\partial^2 \log L}{\partial u_i \partial u_j}\right\} \tag{28}$$

and $g(u) = g\left(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{\lambda},\hat{\sigma}_c^2,\hat{\sigma}_s^2\right)$ is a differentiable function of the parameters u, then the variance of g(u) is asymptotic to

$$[g'(u)]^T V^{-1} g'(u) \tag{29}$$

where $$g'(u) = \left(\frac{\partial g}{\partial u_1},\cdots,\frac{\partial g}{\partial u_6}\right)^T.$$

Thus, in our case, $g(u) = \dfrac{\sigma_s^2}{\lambda^2\sigma_s^2 + \sigma_c^2}$

$$g'(u) = \left[0,0,0,\ -2\lambda\sigma_s^4\left(\lambda^2\sigma_s^2 + \sigma_c^2\right)^{-2}, -\sigma_s^2\left(\lambda^2\sigma_s^2 + \sigma_c^2\right)^{-2}, \frac{1}{\left(\sigma_c^2 + \lambda^2\sigma_s^2\right)} - \frac{\lambda^2\sigma_s^2}{\left(\sigma_c^2 + \lambda^2\sigma_s^2\right)^2}\right] \tag{30}$$

To estimate τ, the observations $\{x_i\}$, $\{y_i\}$, and $\{Z_i\}$ and the estimated parameters $\left(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{\lambda},\hat{\sigma}_c^2, \text{ and } \hat{\sigma}_s^2\right)$ were substituted into the matrix $H = (h_{ij}) = \dfrac{\partial^2 \log L}{\partial u_i \partial u_j}$. Then equation (29) was used to obtain an approximate variance for $\hat{\rho}$.

PRECEDING PAGE BLANK NOT FILMED

APPENDIX D

INSTRUCTIONS TO USDA/ASCS

FOR

- LACIE BLIND SITE SEGMENT INVENTORY
- LACIE BLIND SITE 15-FIELD PERIODIC OBSERVATIONS

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS 77058

AMERICAN REVOLUTION BICENTENNIAL 1776-1976

REPLY TO ATTN OF: TF4-76-11-31

NOV 22 1976

TO: County Executive Director

FROM: Bobby E. Spiers
Production Specialist

SUBJECT: Transmittal of Material for LACIE Segment Inventory

Enclosed is the material to be used to complete the identification of crops in the LACIE segment located in your county. The information requested is essential to the evaluation of the effectiveness of our experiment and we ask that you make the inventory and return the data to us within two weeks.

The instructions explain in detail what needs to be done. Read the instructions and if you have questions, let us know. Thank you for your assistance.

Bobby E Spiers

Bobby E. Spiers

Enclosures
1. Instructions
2. Post Card
3. Return Label
4. Color Infrared Print(s)

FALL

INSTRUCTIONS

FOR

LACIE SEGMENT INVENTORY

TABLE OF CONTENTS

I. Introduction

A. Background

The LACIE (Large Area Crop Inventory Experiment) is an interagency experiment in the use of Landsat (formerly called Earth Resources Technology Satellite) and meteorological data to identify and inventory crop production. Participating agencies include the Department of Agriculture, the National Aeronautics and Space Administration, and the National Oceanic and Atmospheric Administration. Within the Department of Agriculture, participating agencies are the Agricultural Stabilization and Conservation Service, Economic Research Service, Foreign Agricultural Service, Agricultural Research Service, Soil Conservation Service, and Statistical Reporting Service. The overall general objectives of the LACIE are to determine utility and cost effectiveness of satellite and surface derived data sources to monitor large area crop (wheat) production and assess the impact of agricultural and meteorological conditions on production estimates. The utility of the information produced will be evaluated on the basis of its objectivity, timeliness, accuracy, and its expected value for policy and program decision making.

LACIE reports are based on data extracted from 5 x 6 mile segments that have been randomly placed throughout the wheat producing region of the United States. In order to determine our accuracy, it is necessary that we know what is actually in our sample segment. The information requested for the segment that has been identified and forwarded to you

is essential for a successful evaluation of the project. The enclosed color prints have been obtained only over the selected site in your county to support ground data collection. This is the only copy of the print and additional copies are not available for distribution to offices.

B. Authority

The USDA LACIE Project Manager has requested that the Agricultural Stabilization and Conservation Service provide this function and they have accepted the assignment. You should have already gotten an authorization from your State office concerning this task. If you have not, you should contact them at once.

C. Requirements of the ASCS County Office

You are being asked to do the following:

1. Review the set of instructions.
2. Visit the segment location and identify the land uses, even if the segment falls outside your county.
3. Check over your work and return the completed inventory as soon as possible.

II. Data Collection Procedures

A. Supplies

1. Color infrared print or prints.
2. Mylar overlay.
3. Topographic map with segment location.
4. Standard crop key.
5. Crop stage development key.
6. Forms to record plant height, ground cover observations and evaluation comments.

7. Return post card and return mailing tube label.

B. In some cases, all of the segment will not be covered by the photo. Complete the survey for that portion outlined on the photo.

C. Procedures

1. You are required to identify all fields within the segment boundaries using codes as indicated on the attached crop key (see attached LACIE segment classification).
2. Use black or red ball point pen for all coding directly on the mylar.
3. The photos are provided as a base for field pattern and references.
   a. All field identification should be based on actual ground conditions on the day that you visit the segment.
   b. If there are any differences between the photo and the ground, then footnote each field that is different and explain on evaluation form.
   c. Fields that are currently idle crop land should be marked as I/(code) and specifically identify the type of cover, i.e., ST = stubble; CC = winter seeded or volunteer growing crop; RE = disked or harrowed with residue; F = fallow (clean tilled).
4. Use the evaluation form for all comments on any unusual crop condition or practice (irregular, replanting, drought, etc.).

5. If there are any crops in the segment for which there is no code, select an unused symbol and indicate its meaning on the evaluation form.

6. Assess the average wheat crop stages while completing the segment inventory and enter it on the evaluation form upon completion.

III. JSC Contact

A. If there are any problems, contact the person listed below.

B. Review procedures and crop key before going into the field and contact the Johnson Space Center if there are any questions.

Bobby E. Spiers
U.S. Department of Agriculture/ASCS
NASA - Johnson Space Center
Large Area Crop Inventory Experiment
Houston, TX 77058

Phone: Commercial - A/C 713-483-4623
FTS - 525-4623

IV. Due Date and Mailing Procedures

A. Upon receipt of data from the Johnson Space Center, complete the enclosed card and return it to JSC.

B. Field information should be collected within 10 days after receipt of material by your office, if at all possible.

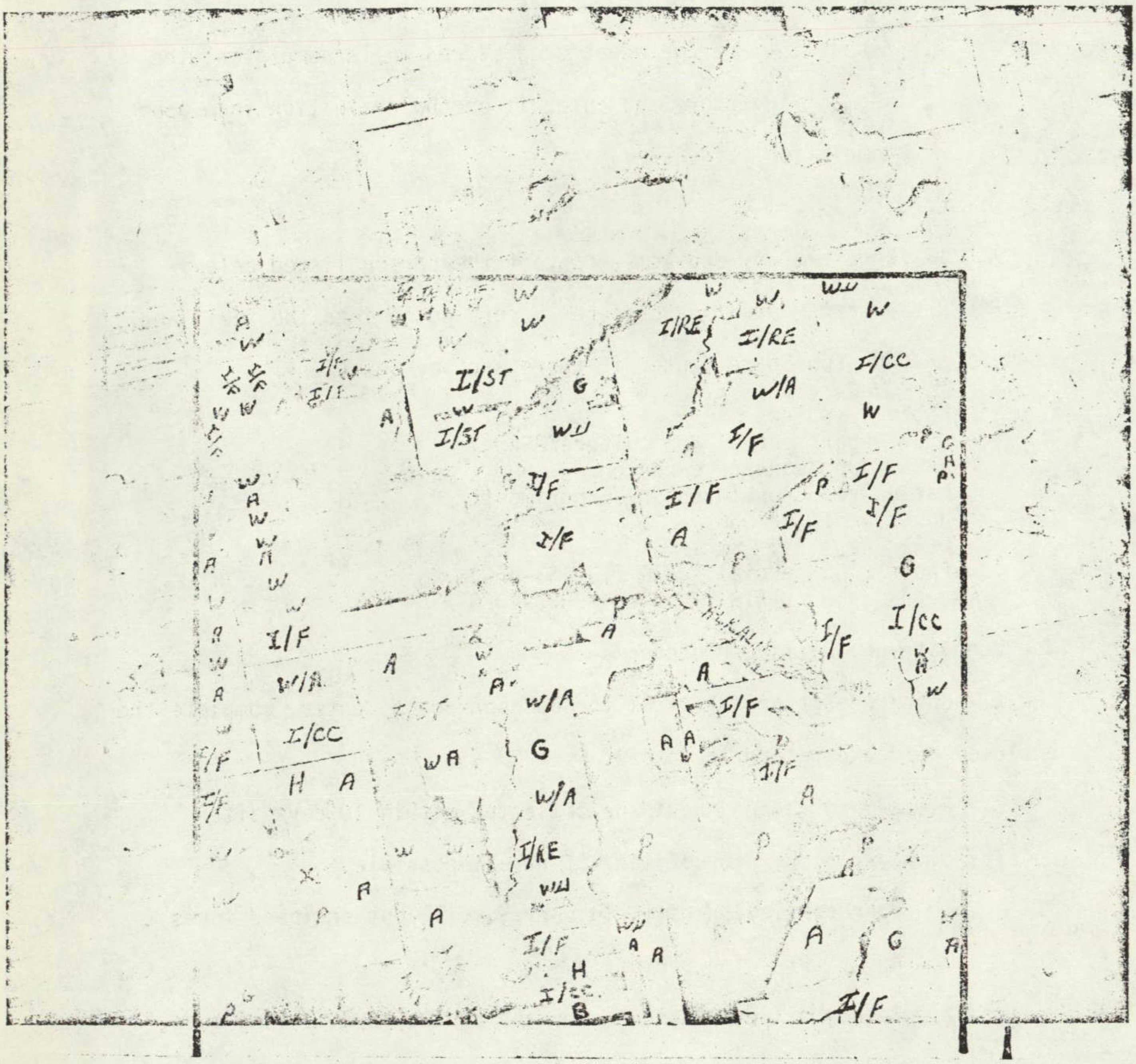
1. Upon completion of field survey, fill out enclosed forms and return with photos.

2. Return all evaluation forms and photos in a mailing tube, using the provided return label.

C. Thank you for your cooperation and effort in assisting LACIE in this vital area of the experiment.

SAMPLE SEGMENT

With Photo and Mylar
(Not to Scale)



Sample: W/A These were winter wheat fields but are currently being plowed under due to the effects of drought.

STANDARD CROP KEY

KEY	CROP TYPE
W	Winter Wheat
SW	Spring Wheat
G	Grass (not cut for hay and no fence)
H	Hay (any visible signs of hay activities)
A	Alfalfa
P	Pasture
C	Corn
SF	Safflower
SU	Sunflower
SG	Sudan Grass
SR	Sorghum
SY	Soybeans
SB	Sugar Beets
FX	Flax
T	Trees
R	Rye
B	Barley
X	Homestead - nonag, lakes, ponds, etc.
BN	Beans
O	Oats
(Crop)/H	Crop has been harvested, 1977 crop
(Crop)/A	Crop has been abandoned; footnote and explain, 1977 crop
I/(Code)	Idle crop land/ST = stubble; cc = growing cover crop; RE = residue; F = fallow (clean tilled)

1. Do not use the code W or SW for any crop other than wheat.

2. If there are crops in segment for which there is no code, select an unused symbol and indicate its use on the evaluation form.

3. Use standard key for all identification.

4. Use ball point pen for all coding on mylar.

CROP STAGE KEY

CROP STAGE KEY	STAGE	DESCRIPTION
1.0	Planted	Seed was put in the ground.
2.0	Emerged	When one leaf per plant is visible.
3.0	Jointed	Defined as when the first node of the stem is visible.
4.0	Heading	Defined as the stage when the base of the rachis (or head) reached the same height as the ligule (or base of the shot leaf).
5.0	Soft Dough	At this stage the crop is starting to turn color. The kernals can be easily deformed when pressed between the fingers, but no "milk" or liquid should exude under such pressure.
6.0	Hard Dough	The kernals readily part from the head. The grain is firm and though it may be dented by pressure of the thumbnail, it is not easily crushed. The characteristic color of the grain has become more distinct. The leaves are brown, dry, and shrunken. Wheat in this stage may be swathed in some areas.
7.0	Harvested or Harvestable	Straw is brittle and dull yellow at this stage. The grain (if not harvested yet) is hard and breaks into fragments when crushed.

EVALUATION FORM

Segment No.: ______ County: ________________ State: ______________

Name: ____ Date: ______________

Man-Hours Required to Complete Survey: __________ Mileage: __________

I. Based on your assessment of the development of wheat in the segment while completing the survey, what is the average wheat stage for the segment? See attached Crop Stage Key. Is the crop development this year in the segment normal, ahead, or behind as compared to previous years? Explain. Enter Crop Stage: ______________________________

II. Comments, footnotes, and additional crop key used:

III. Comments on the effects of drought and/or winterkill:

IV. Comments and recommendations for improving these procedures for future surveys:

Plant Height
and
Percent Ground Cover Observations

Segment Number____________
Observation Date____________

1) During and/or after your segment inventory complete this form and obtain plant height and cover in 15 segment wheat fields.

2) These fields should represent a range from below average to above average stands of wheat for the sample segment only.

3) Select five fields in each category below average, average, and above average based on your judgment of the overall average condition of wheat in the sample segment only.

4) Plant height and percent ground cover should be measured at the same location within the field and should represent the average condition of the field.

5) Numbers should be assigned to fields according to the table below, fields 1-5 below average, 6-10 average, 11-15 above average. Fields can be selected and data collected in any order.

6) Plant height should be measured to the nearest inch.

7) Ground cover should be measured according to the following codes: 1(0-19%), 2(20-39%), 3(40-59%), 4(60-79%), 5(80-100%). This measurement is taken by looking straight down at the ground and estimating the percent of soil that is covered by the crop, and assigning the appropriate code to the field.

Data Collection Requirements:

You are required to select fields and determine which of three categories the field should be assigned to. Select the next available field number in that category. Enter the field from the most accessible corner or turn row. Proceed into the field at least 20 paces to a spot that you feel represents the average condition of the field and proceed to obtain plant height and percent ground cover. Write the field number on the overlay at the exact location that the measurements were obtained and circle the number. Continue with the inventory or field selection.

Below Average			Average			Above Average		
Field No.	Plant Height	Percent Ground Cover Code	Field No.	Plant Height	Percent Ground Cover Code	Field No.	Plant Height	Percent Ground Cover Code
1			6			11		
2			7			12		
3			8			13		
4			9			14		
5			10			15		

APPENDIX E

FORMS USED IN TAKING INVENTORIES OF LACIE INTENSIVE TEST SITES

APPENDIX E

FORMS USED IN TAKING INVENTORIES OF LACIE INTENSIVE TEST SITES

The following forms are used to record ground observations in the LACIE ITS's:

Figure	Form
E-1	Sample Ground Truth Inventory Form
E-2	Sample Ground Truth Periodic Observation Form
E-3	Sample form for Ground Truth Data Collection System Rainfall Measurements
E-4	Sample Yield Form

TEST SITE # 25 (SHELBY, IND.)
INVENTORY OBSERVATION

GROUND TRUTH INVENTORY FORM

DATA RECORDED BETWEEN 11/1/75 AND 11/10/75

LAND USE CROP CODES

100 SPRING WHEAT (GEN) CODE LIST ATTACHED	211 PARAGON	306 KINSEY	410 WANSER	423 CAPROCK	606 REDWOOD FLAX
200 BARLEY (GEN)	212 HERTA	307 HARMON	411 MCCALL	424 EAGLE SCOUT	607 CORN
201 CARRER	213 CONQUEST	308 MISSION	412 SHEVOR	425 TRAPPER	610 BUCKWHEAT
202 VANGUARD	214 WISCONSIN	400 WINTER WHEAT (GEN)	413 GAINES	426 CHEYENNE	611 SUNFLOWERS
203 BONANZA	215 HECTOR	401 DARK HARD NORTHERN	414 NUGAINES	427 WINALTA	612 DRY BEANS
204 DICKSON	216 HIPANA	402 SCOUT	415 DRUCHAMP	428 WINOKA	613 LENTILS
205 PRIMUS	217 MORAVIAN	403 PARKER	416 MORO	500 GRASSES/PASTURE ETC.	614 DRY PEAS
206 SHABET	300 OATS (GEN)	404 EAGLE	417 BURT	600 OTHER CROPS (GEN)	615 SUGAR BEETS
207 PIRELINE	301 RUSSELL	405 APACHE	418 BRIDGER	601 RAPESEED	616 GRAIN SORGHUM
208 BETZES	302 BASIN	406 TRIUMPH	419 ARTHUR	602 RYE	617 SOYBEANS
209 STEPTOE	303 KELSEY	407 SATANTA	420 TASCOSA	603 MUSTARD	618 COTTON
210 FERGUS	304 RODNEY	408 CENTURK	421 STURDY	604 FLAX (GEN)	700 SUMMER FALLOW
	305 LODI	409 BISON	422 CONCHO	605 NORALTA FLAX	800 NON-AGRICULTURE
					900 UNKNOWN CROPS

MAP REFERENCE # OF FIELD	ACREAGE	LAND USE CROP CODE	IRRIGATED	FERTILIZED	PLANTING DATE MONTH DAY YEAR	COMMENTS
16	15.0	402	Y N	Y N	10 / 11 / 75	135# 8-26-26
23	13.3	401	Y N	Y N	9 / 27 / 75	135# 8-26-26
28	5.0	417	Y N	Y N	10 / 5 / 75	200# 6-24-24
35	16.5	402	Y N	Y N	10 / 2 / 75	200# 12-12-12 in spring
61	18.7	417	Y N	Y N	10 / 10 / 75	200# 6-24-24 30# N in spring
76	3.7	417	Y N	Y N	10 / 13 / 75	200# 6-24-24 30# N in spring
97	13.8	417	Y N	Y N	10 / 13 / 75	200# 6-24-24 30# N in spring
107	8.7	417	Y N	Y N	10 / 11 / 75	200# 6-24-24 30# N in spring
108	9.8	417	Y N	Y N	10 / 11 / 75	200# 6-24-24 30# N in spring
118	12.6	417	Y N	Y N	10 / 5 / 75	200# 5-20-20
137	10.8	417	Y N	Y N	10 / 5 / 75	200# 6-24-24
140	24.8	417	Y N	Y N	10 / 5 / 75	200# 6-24-24
147	3.8	417	Y N	Y N	10 / 4 / 75	200# 7-26-26 35# N in spring
157	23.0	417	Y N	Y N	10 / 4 / 75	200# 7-26-26 35# N in spring
162	14.1	402	Y N	Y N	10 / 1 / 75	200# 7-26-26 35# N in spring
~~173~~		~~7 0~~	~~Y N~~	~~Y N~~	~~7 / 30 / 75~~	~~[illegible]~~
190	22.6	417	Y N	Y N	9 / 30 / 75	200# 5-20-20
193	27.3	417	Y N	Y N	10 / 13 / 75	200# 6-24-24
194	7.0	417	Y N	Y N	10 / 11 / 75	200# 6-24-24 30# N in spring
205	14.5	417	Y N	Y N	10 / 11 / 75	200# 6-24-24 30# N in spring
224	15.0	417	Y N	Y N	10 / 12 / 75	Will apply N + fert in spring [illegible]
226	8.[illegible]	417	Y N	Y N	10 / 10 / 75	[illegible]
231	12.0	402	Y N	Y N	10 / 1 / 75	Unsure 200# 6-24-24
254	10.0	417	Y N	Y N	10 / 1 / 75	
313	16.9	419	Y N	Y N	10 / 13 / 75	225# 6-24-24

Figure E-1.– Sample Ground Truth Inventory Form.

GROUND TRUTH PERIODIC OBSERVATION FORM

TEST SITE # 25 (SHELBY, INDIANA)
OBSERVATION 01

LANDSAT PASS DATE 4 / 27 /76 (MONTH / DAY)
OBSERVATION DATE 4 / 31 /76
RAINFALL SINCE LAST OBSERVATION ___ IN.

LAND USE CODES	GROWTH STAGES	GROUND COVER (%)	SURFACE MOISTURE CONDITIONS	FIELD OPERATIONS	GROWTH/YIELD DETRACTANTS	STAND QUALITY
100-SPRING WHEAT 200 BARLEY 300-OATS 400-WINTER WHEAT 500-GRASSES/PASTURE 600-OTHER CROPS 601-RAPESEED 602-RYE 604-FLAX 607-CORN 617-SOYBEANS 618-COTTON 700-SUMMER FALLOW 900-UNKNOWN CROPS	01-NOT PLANTED 02-PLANTED NO EMERGENCE 03-EMERGENCE 04-TILLERING, PREBOOT, PREBUD 05-BOOTED OR BUDDED 06-BEGINNING TO HEAD OR FLOWER 07-FULLY HEADED OR FLOWERED 08-BEGINNING TO RIPEN 09-RIPE MATURE 10-HARVESTED 11-DOES NOT APPLY	1- 0- 19 2-20- 39 3-40- 59 4-60- 79 5-80-100	1-DRY 2-DAMP 3 WET 4-STANDING WATER WEED GROWTH 1-NEGLIGIBLE 2-SLIGHT 3-MODERATE 4-HEAVY	01-BARE GROUND 02-BARE DISKED/CULTIVATED 03-BARE PLOWED 04-BARESEEDED 05-STANDING STUBBLE 06 STUBBLE DISKED/CULTIVATED 07-STUBBLE PLOWED 08-STUBBLE SEEDED 09-BURNED 10-GRAZED 11-WINDROWED OR SWATHED 12-MOWED OR COMBINED 13 STACKED OR BALED 14-OTHER	01-SALINITY 02-INSECTS 03-DISEASE 04-DROUGHT 05-MOISTURE 06-WIND 07 HAIL 08-FROST 09-BIRDS 10-POT HOLES 11-UNEVEN STAND 12-WEEDS 13-WINTERKILL 14-LODGING 15-OTHER	1-POOR 2-BELOW AVERAGE 3-AVERAGE 4-ABOVE AVERAGE 5 EXCELLENT 6-DOES NOT APPLY

FIELD NO.	ACREAGE	LAND USE CODE	GROWTH STAGE (CIRCLE ONE)	GROUND COVER (CIRCLE ONE)	PLANT HEIGHT (INCHES)	SURFACE MOISTURE (CIRCLE ONE)	WEED GROWTH (CIRCLE ONE)	FIELD OPERATIONS (CIRCLE ONE)	GROWTH/YIELD DETRACTANTS (CIRCLE ONE)	STAND QUALITY RATING (CIRCLE ONE)	COMMENTS
903	17.0	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	16	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
86	31.5	607	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	0	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
52	48.2	607	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	0	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
140	24.8	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	14	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
92	19.0	617	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	0	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
16	15.0	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	16	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
64	38.3	607	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	0	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
806	24.0	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	12	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
35	16.6	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	14	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
75	27.3	607	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	0	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
909	19.8	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	16	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
69	18.7	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	14	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
224	16.0	400	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	16	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N
163	19.2	300	01 02 03 04 05 06 07 08 09 10 11	1 2 3 4 5	3	1 2 3 4	1 2 3 4	01 02 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 6	Y N

Figure E-2.— Sample Ground Truth Periodic Observation Form.

ORIGINAL PAGE IS OF POOR QUALITY

TEST SITE # 30 (SALINE, KANSAS)

GROUND TRUTH DATA COLLECTION SYSTEM

RAINFALL MEASUREMENTS

FIELD #	DATE OBSERVED MONTH DAY	DATE RAINED MONTH DAY	RAINFALL	FIELD #	DATE OBSERVED MONTH DAY	DATE RAINED MONTH DAY	RAINFALL
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.
______	__/__/76	__/__/76	______IN.	______	__/__/76	__/__/76	______IN.

No Measurable rain this period

Figure E-3.— Sample form for Ground Truth Data Collection System Rainfall Measurements.

TEST SITE: 37 (FRANKLIN, IDAHO) 1975 YIELD FORM DATE RECORDED BETWEEN 09/29/75 AND 12/30/75 P.1 of 2

ORIGINAL PAGE IS OF POOR QUALITY

FIELD #	LAND USE CODE	HARVEST DATE MO/DAY/YR	ASCS EST	FARMER EST	SAMPLE WT. BU/ACRE	FCIC FCIC EST	STAND QUALITY RATING	COMMENT
7	100	9-6-75	20	10	Y N	Y N	1 2 3 4 5	Y N
18	100	9-6-75	100	70	Y N	Y N	1 2 3 4 5	Y N
19	100	9-6-75	100	70	Y N	Y N	1 2 3 4 5	Y N
46	400	9-6-75	15	13	Y N	Y N	1 2 3 4 5	Y N
98	100	9-15-75	90	65	Y N	Y N	1 2 3 4 5	Y N
181	400	8-19-75	20	18	Y N	Y N	1 2 3 4 5	Y N
132	413	9-6-75	110	100	Y N	Y N	1 2 3 4 5	Y N
247	400	8-19-75	25	22	Y N	Y N	1 2 3 4 5	Y N
229	100	9-6-75	100	105	Y N	Y N	1 2 3 4 5	Y N
338	100	9-6-75		80	Y N	Y N	1 2 3 4 5	Y N
165	400	9-6-75	20	18	Y N	Y N	1 2 3 4 5	Y N
164	400	9-6-75		18	Y N	Y N	1 2 3 4 5	Y N
31	400	9-6-75	35	28	Y N	Y N	1 2 3 4 5	Y N
48	400	9-6-75	12		Y N	Y N	1 2 3 4 5	Y N
375	413	9-6-75	100	97	Y N	Y N	1 2 3 4 5	Y N
170	100	9-6-75		93	Y N	Y N	1 2 3 4 5	Y N
352	100	9-6-75		45	Y N	Y N	1 2 3 4 5	Y N
368	400	8-19-75		14	Y N	Y N	1 2 3 4 5	Y N
414	100	9-6-75		105	Y N	Y N	1 2 3 4 5	Y N
441	414	-/-/75		105	Y N	Y N	1 2 3 4 5	Y N
		-/-/75			Y N	Y N	1 2 3 4 5	Y N
		-/-/75			Y N	Y N	1 2 3 4 5	Y N

Figure E-4.— Sample Yield Form.

APPENDIX F

CROP GROWTH STAGE REPORTING FORMS

Check one:

☐ Spring Wheat

☐ Winter Wheat

Crop Year______

GROWTH STAGE DATES FOR 10% DEVELOPMENT

MONTH AND DAY OF MONTH

ITS SEG. NO.	CRD #	PLANTING DATE[1]	EMERGENCE DATE[1]	JOINTING DATE[2]	HEADING DATE[2]	SOFT DOUGH DATE[3]	RIPE DATE[4]	HARVEST DATE[5]

[1] Date at which 10% of fields in CRD were planted or emerged, respectively.

[2] Date at which 10% of fields in CRD had begun to joint or head, respectively.

[3] Date at which 10% of fields in CRD had begun to enter soft dough stage (turning color to greenish-yellow to yellow).

[4] Date at which 10% of fields in CRD are ripe (hard dough) stage or when they were swathed. (Indicated swathed if applicable).

[5] Date at which 10% of fields in CRD have been harvested either as standing grain or out of swath.

Check one:

☐ Spring Wheat

☐ Winter Wheat

Crop Year______

GROWTH STAGE DATES FOR 50% DEVELOPMENT

MONTH AND DAY OF MONTH

ITS SEG. NO.	CRD #	PLANTING DATE[1]	EMERGENCE DATE[1]	JOINTING DATE[2]	HEADING DATE[2]	SOFT DOUGH DATE[3]	RIPE DATE[4]	HARVEST DATE[5]

[1] Date at which 50% of fields in CRD were planted or emerged, respectively.

[2] Date at which 50% of fields in CRD had begun to joint or head, respectively.

[3] Date at which 50% of fields in CRD had begun to enter soft dough stage (turning color to greenish-yellow to yellow).

[4] Date at which 50% of fields in CRD are ripe (hard dough) stage or when they were swathed. (Indicated swathed if applicable).

[5] Date at which 50% of fields in CRD have been harvested either as standing grain or out of swath.

Check one:

☐ Spring Wheat

☐ Winter Wheat

Crop Year______

GROWTH STAGE DATES FOR 90% DEVELOPMENT

MONTH AND DAY OF MONTH

ITS SEG. NO.	CRD #	PLANTING DATE[1]	EMERGENCE DATE[1]	JOINTING DATE[2]	HEADING DATE[2]	SOFT DOUGH DATE[3]	RIPE DATE[4]	HARVEST DATE[5]

[1] Date at which 90% of fields in CRD were planted or emerged, respectively.

[2] Date at which 90% of fields in CRD had begun to joint or head, respectively.

[3] Date at which 90% of fields in CRD had begun to enter soft dough stage (turning color to greenish-yellow to yellow).

[4] Date at which 90% of fields in CRD are ripe (hard dough) stage or when they were swathed. (Indicated swathed if applicable).

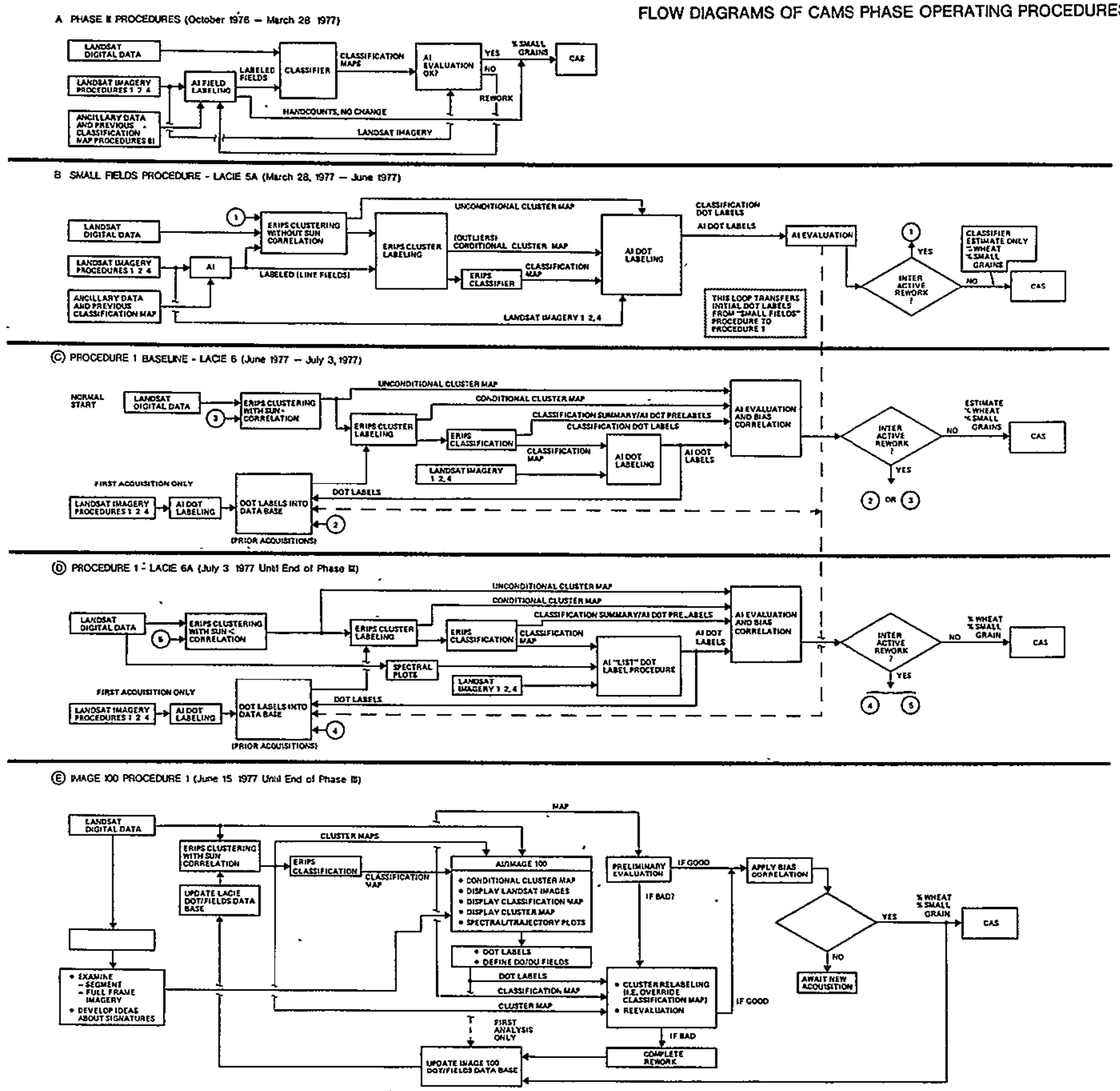
[5] Date at which 90% of fields in CRD have been harvested either as standing grain or out of swath.

PRECEDING PAGE BLANK NOT FILMED

APPENDIX G

PHASE III CAMS OPERATIONAL PROCEDURES FLOW DIAGRAMS

FLOW DIAGRAMS OF CAMS PHASE OPERATING PROCEDURES



APPENDIX H

PHASE III AA PIXEL-LEVEL DATA PROCESSING SYSTEM

APPENDIX H

PHASE III AA PIXEL-LEVEL DATA PROCESSING SYSTEM

In Phase III, AA will conduct pixel-level investigations for the first time. These are designed to monitor the performance of the various components of the LACIE system (analysts, clustering algorithm, and classification algorithm) and to investigate the sources of error in these elements. These investigations are described in section 6.2.5.2. This appendix describes the data processing system that has been designed to support the pixel-level investigations. In addition, this system automatically produces some data (such as segment wheat proportions) that are used in segment-level investigations.

The system involves the processing of ground-truth data, CAMS classification maps, CAMS cluster maps, and CAMS analyst-interpreter-selected dot labels. A flow diagram of this AA data processing system is presented in figure H-1. The Phase III AA processing described in this figure includes four major steps.

1. CAMS DTRM tape preprocessing
2. AA ground-truth preprocessing
3. CAMS analyst dot label preprocessing
4. AA classification/label accuracy processing

The first three steps are necessary to prepare for the fourth one. This last processing step is done using the computer programs SPATL and SPECTL on the PDP 11/45.

Since only about 10 percent of the segments on the DTRM tapes are U.S. blind sites, a preprocessing program is required. This

program will strip the blind site data from the DTRM tape and put it onto another tape containing only blind site data.

The DTRM tape contains all of the classification maps and conditional and unconditional cluster maps for all segments processed by CAMS, beginning with the segments processed using the small fields procedures. Since the identity of the blind sites is to remain unknown to the CAMS analysts, AA will provide blind site identification cards as input for the DTRM tape-searching program. When data for a blind site are located on the DTRM tape, they will be copied onto the AA blind site class/cluster map tape (the first step in the processing task).

The second step is the conversion of ground-truth maps into ground-truth tapes suitable for subsequent processing. This initially involves the use of the Dell Foster equipment which produces field vertices cards. These cards are converted to a tape by the program BTAPE to produce a tape with a format similar to the tape created by the Bendix 100 Interactive Drafting System. This format contains the x and y coordinates of the vertices of the agricultural fields from the aerial photographs of the blind sites. The resultant product is a nine-track computer-compatible tape (CCT) containing a field identification and field coordinates.

The Dell Foster equipment will be used for early season processing until the Bendix 100 software package for the Interactive Drafting System becomes available and is checked out. The schedule for the beginning of ground-truth map processing using the Bendix 100 is June 15, 1977. The Bendix 100 tapes will be processed by the programs PHASE1 and PHASE2 to produce a Universal format ground-truth tape containing AA crop labels at the subpixel level, registered to the Landsat coordinates. This AA ground-truth tape will be coded as per the AA codes presented in

table 6-2. The ground-truth tape is the product of the second step in the AA processing system.

The third step requires SF3 to acquire the SF4 CAMS pixel-label sheets which contain the analyst dot labels and classification dot labels prepared during the Phase III operational processing. These sheets were initially used in operations when the small fields procedures were implemented and will be used throughout the remainder of Phase III. From these sheets, SF3 will keypunch dot label cards to be used as input to the AA accuracy calculations (step 4). The AI dot label card deck or tape is the output of step 3. As Procedure 1 implementation progresses, the CAMS/CAS interface tape may be used to obtain the analyst-interpreter dot labels instead of the analyst-interpreter dot label cards.

The AA classification/labeling processing element utilizes programs SPATL and SPECTL to compare the class/cluster map tape, the ground-truth tape, and the analyst-interpreter dot labels. This comparison is done at the pixel and segment levels, which generate data for subsequent AA evaluation. For additional information regarding this processing step, refer to technical memorandum LEC-10620 (ref. 4), entitled "Requirements Document For Phase III LACIE Accuracy Assessment."

Later in Phase III, an AA data file will be used as indicated in figure H-1. The PDP 11/45 will be used to store the three basic inputs on the disk (class/cluster maps, ground truth, and analyst dot labels). The fourth-element programs will use this disk as a data base.

Figure H-1 indicates the processing responsibilities of SF3 and SF12. The responsibilities of SF3 are contained within the large dashed lines, with the remaining elements being the responsibility of SF12.

In order to assure that the results from the programs SPECTL and SPATL are correct, some of the data should be processed interactively on the IMAGE 100 system. It is anticipated that about one segment per week or about 5 percent of the blind sites should be processed in this manner. About 4 hours of IMAGE 100 time per segment will be required.

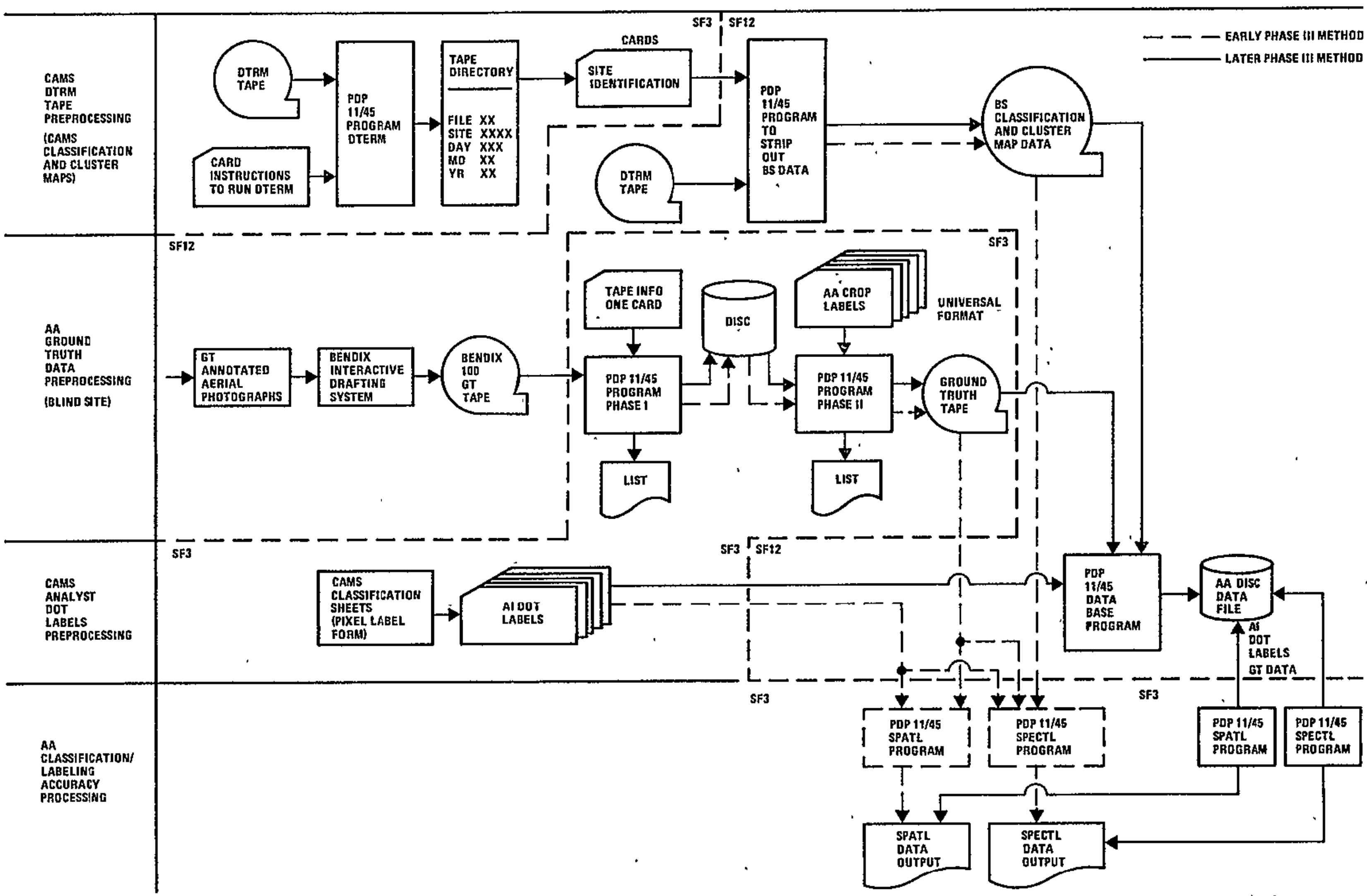


Figure H-1.– Phase III Accuracy Assessment data processing system flow diagram.

PRECEDING PAGE BLANK NOT FILMED

APPENDIX I

ABBREVIATIONS, ACRONYMS, AND SYMBOLS

APPENDIX I

ABBREVIATIONS, ACRONYMS, AND SYMBOLS

AA	Accuracy Assessment.
AA-01	LACIE Phase II AA Report for February 1976.
ACC	adjustable crop calendar.
Agromet	agricultural/meteorological.
AI	analyst-interpreter.
Biowindow or biophase	biological window – a Landsat data acquisition period that is related to the biostages of wheat development. The LACIE approach is based on the judgment that wheat can be separated adequately from other crops by analysis of up to four acquisitions of Landsat data during the growing season. The biowindow may be updated if there is a significant lag or advancement in the current crop calendar. The sequence chosen includes acquisitions during the following biowindows: 1. Crop establishment – from 50 percent tillering to 50 percent jointing (biostage 2.3 to 3.0). 2. Green – from 50 percent jointing to 50 percent heading (biostage 3.1 to 4.0). 3. Heading – from 50 percent heading to 50 percent soft dough (biostage 4.1 to 5.0). 4. Mature – from 50 percent soft dough to 50 percent harvest (biostage 5.1 to 6.0).
Biostage	biological stage – the specific stage of development of a crop which can be recognized by a major change in plant structure; i.e., emergence after germination, jointing, heading, soft dough, ripening, and harvest, which are represented by integers on the Robertson Biometeorological Time Scale.

Blind site	a LACIE sample segment that is a part of the LACIE operational random sample set of segments used in the LACIE aggregations. The blind sites are *not* identified to analysts in order to ensure normal processing of the blind site segments. However, in Canada, where only the blind site segments are being processed, the analyst obviously knows their location.
BMTS	Biometeorological Time Scale.
CAMS	Classification and Mensuration Subsystem.
CAR	CAS Annual Report
CAS	Crop Assessment Subsystem.
CCEA	Center for Climatic and Environmental Assessment — an organization of the National Oceanic and Atmospheric Administration, Columbia, Missouri.
Classification	in computer-aided analysis of remotely sensed data, the process of assigning data points to specified classes by a testing process in which the spectral properties of each unknown data point are compared with spectral properties typical of the subject being classified.
Classification error	a measure of the degree to which the LACIE CAMS overestimates or underestimates the wheat acreage in one or more LACIE samples.
CMR	CAS Monthly Report.
COM	classification on microfilm.
CRD	Crop Reporting District — a geographical area used by the U.S. Department of Agriculture for the collection and reporting of agricultural information; each district consists of several counties.
Crop calendar	a calendar depicting the biostages of the major crop types within a specified region during a calendar year.

Crop calendar adjustment	an adjustment made to the normal crop calendar on the basis of current meteorological data.
CUR	CAS Unscheduled Report.
CV	coefficient of variation (standard deviation divided by the mean).
DAPTS	Data Acquisition, Preprocessing, and Transmission Subsystem.
DO/DU	designated other/designated unidentifiable.
DOY	day of the year
First-order error source evaluations	evaluations to identify the major components of error in LACIE Phase III production estimates; these components can be statistically estimated with techniques described in section 6.2.4 of the report.
Group 2 segment	LACIE segment in a county that historically produces small quantities of wheat/small grains; samples are allocated with probability proportional to size.
Group 3 segment	LACIE segment in a county that historically produces very small quantities of wheat/small grains; estimates are based on the changes in acreage of group 1 and 2 segments from year to year.
IE	Information Evaluation.
IMR	IE Monthly Report.
ISRRS	Information, Storage, Retrieval, and Reformatting Subsystem
ITS	intensive test site - a LACIE segment in the United States or Canada over which detailed crop information is collected by using ground and airborne observation equipment; these ITS segments are a separate set of segments from the operation LACIE segments used in the aggregations.
JSC	Lyndon B. Johnson Space Center of NASA, Houston, Texas.
LACIE	Large Area Crop Inventory Experiment.

Landsat	Land Satellite – formerly called ERTS (Earth Resources Technology Satellite); operates in a circular, Sun-synchronous, near-polar orbit of the Earth at an altitude of approximately 915 kilometers; orbits the Earth about 14 times a day and views the same scene approximately every 18 days.
LEC	Lockheed Electronics Company, Inc.
LPDL	LACIE Physical Data Library.
LPP	LACIE Performance Predictor.
MSS	Multispectral Scanner System or multispectral scanner – the remote sensing instrument on Landsat that measures reflected sunlight on various spectral bands or wavelengths.
Multitemporal analysis	analysis of data sets over the same area acquired at different times during the growing season.
NASA	National Aeronautics and Space Administration.
NOAA	National Oceanic and Atmospheric Administration.
90-90 criterion	criterion that the LACIE U.S. Great Plains production estimate is 90 percent accurate, at harvest, 90 percent of the time (in comparison with the true value).
PCC	probability of correct classification.
PFC	production film converter.
PPS	probability proportional to size.
RD	relative difference.
RTEB	Research, Test, and Evaluation Branch.
Sample segment	a 5- by 6-nautical-mile area selected by stratified random sampling; information is recorded by the MSS and transformed into computer-compatible tapes and film products.
Sampling error	a measure of the degree to which the estimated wheat acreage in the LACIE sample segments does not represent the wheat acreage contained in the survey region being sampled.

Second-order error source evaluations	evaluations to further qualitatively break down the identified first-order components into factors that can be related to causal elements in LACIE methods and procedures.
USDA	U.S. Department of Agriculture.
USDA/ASCS	USDA Agricultural Stabilization and Conservation Service.
USDA/FAS	USDA Foreign Agricultural Service.
USDA/SRS	USDA Statistical Reporting Service.
U.S. Great Plains	an area encompassing the nine states of Colorado, Kansas, Minnesota, Montana, Nebraska, North and South Dakota, Oklahoma, and Texas; it is divided geographically into (1) the U.S. southern Great Plains, which includes Colorado, Kansas, Nebraska, Oklahoma, and Texas, and (2) the U.S. northern Great Plains, which includes Minnesota, Montana, and North and South Dakota.
WMO	World Meteorological Organization.
YES	Yield Estimation Subsystem.

SYMBOLS:

$\hat{P}$	wheat/small-grain proportion estimate.
P_W	proportion of wheat harvested.
σ	standard deviation.
$CV(\hat{W})$	CV for production estimate.
P_{GT}	proportion of wheat/small grains based on identification of each field in the blind site or ITS by USDA/ASCS personnel.

APPENDIX J

REFERENCES

APPENDIX J

REFERENCES

1. Large Area Crop Inventory Experiment (LACIE) Crop Assessment Subsystem (CAS) Requirements. Vol. IV, LACIE-C00200 (rev. B), JSC-11329, July 1976.

2. Large Area Crop Inventory Experiment (LACIE) Yield Estimation Subsystem (YES) Requirements. Vol. III, LACIE-C00200 (rev. A), JSC-11340, Sept. 1976.

3. Large Area Crop Inventory Experiment Crop Assessment Subsystem Software Requirements Document. JSC-10009, Nov. 25, 1975.